

78041

STIC-Biotech/ChemLib

195022

M9

From: Bunner, Bridget
Sent: Monday, July 10, 2006 10:14 AM
To: STIC-Biotech/ChemLib
Subject: sequence search request

Hi! I'd like to request a sequence search for case 10/724,806:

1. the nucleic acid sequence that encodes SEQ ID NO:4
2. the nucleic acid sequence of SEQ ID NO: 3

Thanks!

Bridget Bunner

Art Unit 1647
Rem 4C65
(571) 272-0881
mailbox 4C70

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:29 ; Search time 9689 Seconds
(without alignments)
11503.804 Million cell updates/sec

Title: US-10-724-806-3
Perfect score: 1743
Sequence: 1 atgccttcacatgctggaag.....ctgaagataattacaatga 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 segs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_srs:*
8: gb_sy:*
9: gb_un:*
10: gb_va:*
11: gb_ov:*
12: gb_hcg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1743	100.0	1743	6	AF276872 Mus muscu
2	1731.8	99.4	4715	6	BC065089 Mus muscu
3	1730.2	99.3	1743	2	BD012720 High-affi
4	1730.2	99.3	1743	2	BD012720 High-affi
5	1730.2	99.3	1743	2	BD012720 High-affi
6	1723.8	98.9	1743	6	MMU01467 Mus muscu
7	1616.6	92.7	1743	2	BD012718 High-affi
8	1616.6	92.7	1743	2	E49870 High-affi
9	1616.6	92.7	4904	6	AB030947 Rattus no
10	1378.2	79.1	1743	2	AR268949 Sequence
11	1375	78.9	1743	2	BD012719 High-affi
12	1375	78.9	1743	2	CQ722844 Sequence
13	1375	78.9	1743	5	E49871 High-affi
14	1375	78.9	1743	5	AF276871 Homo sapi
15	1375	78.9	1813	5	HS401466 Homo sapi
16	1375	78.9	2583	5	BC111525 Homo sapi
17	1375	78.9	5158	5	AB043997 Homo sapi
18	1297.4	74.4	2541	5	BC111524 Homo sapi

19	1243.8	71.4	2112	14	BC112794 Bos tauru
20	841.4	48.3	2528	11	TMA420808 Torpedo m
21	699	40.1	1132	11	GA511267 Gallus ga
22	629.2	36.1	155210	6	AC102873 Mus muscu
23	629.2	36.1	257456	12	CT571242 Mus muscu
24	600.4	34.4	232792	12	AC106657 Rattus no
25	504.4	28.9	159918	12	AC182392 Pan trogl
26	502.8	26.8	2239	5	HS408384 Homo sapi
27	502.8	26.8	190043	5	AC009963 Homo sapi
28	437.8	25.1	3326	13	AY011119 Limulus p
29	421.2	24.2	3255	13	AY047521 Drosophila
30	372.6	21.4	2049	13	AY629593 Trichoplui
31	356.4	20.4	386	2	AX080449 Sequence
32	355.8	20.4	1731	2	BD012717 High-affi
33	355.8	20.4	1731	2	E49869 High-affi
34	355.8	20.4	1985	13	AB030946 Caenorhab
35	321	18.4	329	2	AX956120 Sequence
36	278.8	16.0	1729	2	C0613836 Sequence
37	257.8	14.8	172639	12	CR925787 Danio rer
38	240.4	13.8	212798	11	BX901972 Zebrafish
39	237.6	13.6	1461	2	AX432086 Sequence
40	237.6	13.6	110000	15	AB017333_31 Continuation (32 o
41	237.6	13.6	110000	15	CP000002_31 Continuation (32 o
42	221.4	12.7	160000	11	AC174718 Gasterost
43	178	10.2	4223	2	C0613835 Sequence
44	178	10.2	186989	13	AC007812 Drosophila
45	178	10.2	189116	13	AC009395 Drosophila

ALIGNMENTS

RESULT 1
LOCUS AF276872 1743 bp mRNA linear ROD 28-FEB-2001
DEFINITION Mus musculus sodium and chloride-dependent high-affinity choline transporter mRNA, complete cds.
ACCESSION AF276872
VERSION AF276872.2 GI:13162669
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1743) Apparsundaram,S., Ferguson,S.M. and Blakely,R.D. Molecular cloning and characterization of human and murine high-affinity choline transporters
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1743) Apparsundaram,S., Ferguson,S.M. and Blakely,R.D.
TITLES Direct Submission
JOURNAL Submitted (09-JUN-2000) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at Pierce, Nashville, TN 37232-6420, USA
REFERENCE 3 (bases 1 to 1743) Apparsundaram,S., Ferguson,S.M. and Blakely,R.D. Direct Submission
TITLES Submitted (28-FEB-2001) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at Pierce, Nashville, TN 37232-6420, USA
REMARK COMMENT On Feb 28 2001 this sequence version replaced gi:11527247.
FEATURES
Location/Qualifiers
1..1743
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..1743
/codon_start=1
/product="sodium and chloride-dependent high-affinity choline transporter"

```

/protein_id="AA036945.2"
/db_xref="GI:13162670"
/translation="MPHVGGLVAIIILFYLLFLVGIWAMAKTKNSGNPERSEAIIVY
GGRIIGLVGGFTMTQVWGGYINGTAEVYVPGGGALMAHAIPGYSLTSLGGF
AKPMKSGVYMLDPFQIIVYQKGGKGIIFPALMGSPMAAIIIFALGATISLIIVD
NIVISVIALSVYITLVGGLVSAVATYDVLQLECIPIGMISFVPSLSPAVTIDGFI
VAHAKYQSPMLGTIESVYETWMLDNLMLTGGIPQAFQGVYVSSSAPTAQVLSF
AAGCLVWALPAITGIGASTDMNQTGAIGYDPKPKKEADNLIIFLYVQICPVYIS
FGLVASAAWWSADSSILSASSPANIYQLSFRGNADKIIWMMIYLVFPAAS
TAMALLRTVYVGLWSSLDYVLIIFPOLCLYFIGNTYVAAGVYIFGLFIRITGG
EPLVYLVAVLIPSYEDYDKNGIYNORPEKTSMTYSPTNIQSVSLAKYIFESGTE
KLVYLVAVYVAPLHSEEMDKTILVLRNENTKLNELAVKRGQSLTISSTPKKALDDV
DSPEGGGCTEDNLO"

```

ORIGIN

Query Match	100.0%;	Score 1743;	DB 6;	Length 1743;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1743; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCGCTTCCATGTCGGAAGGACGTGGTACCTATATATCTTCTTACTCTCTTAAATTTCTG	60
Db	1	ATGCGCTTCCATGTCGGAAGGACGTGGTACCTATATATCTTCTTACTCTCTTAAATTTCTG	60
QY	61	GTTGGAAATATGGGCTGCATGGAACCAAAAAACAGCGGCAACCCAGAAAGCGCAGTGA	120
Db	61	GTTGGAAATATGGGCTGCATGGAACCAAAAAACAGCGGCAACCCAGAAAGCGCAGTGA	120
QY	121	GCCATCATAGTCGGGGGCGGTGACATTGGTTGTTGGTGGTTTAAACATGACGCC	180
Db	121	GCCATCATAGTCGGGGGCGGTGACATTGGTTGTTGGTGGTTTAAACATGACGCC	180
QY	181	ACCGGGGTGAGAGAGGCTACATCAATGGGACAGAGAAAGCAGTATGGGCGAGTGT	240
Db	181	ACCGGGGTGAGAGAGGCTACATCAATGGGACAGAGAAAGCAGTATGGGCGAGTGT	240
QY	241	GGTCTAGCTTGGGCTCATGCAACCCATTGGATATTTCTGAGCTATATTTAATTAAGTGTG	300
Db	241	GGTCTAGCTTGGGCTCATGCAACCCATTGGATATTTCTGAGCTATATTTAATTAAGTGTG	300
QY	301	TTTTTTGCGAAACCTATGCGTTCAGAGGAAATGTGACTATGTTAGACCCATTCAACAG	360
Db	301	TTTTTTGCGAAACCTATGCGTTCAGAGGAAATGTGACTATGTTAGACCCATTCAACAG	360
QY	361	ATCTATGGAAGCGCATGGGTGGGCTCTTATTCCTGCACTATATGGAGAAATGTC	420
Db	361	ATCTATGGAAGCGCATGGGTGGGCTCTTATTCCTGCACTATATGGAGAAATGTC	420
QY	421	TGGGCTGAGCAATTTCTGTCATTTAGGGGCACTATCAGGAGTCAATGATGAT	480
Db	421	TGGGCTGAGCAATTTCTGTCATTTAGGGGCACTATCAGGAGTCAATGATGATGAT	480
QY	481	GTTGAACATATCGGTATGTTCTTGGCACTATTTGCCATTTCTTAAACCTATAGTGGTGG	540
Db	481	GTTGAACATATCGGTATGTTCTTGGCACTATTTGCCATTTCTTAAACCTATAGTGGTGG	540
QY	541	CTTAATCTCTGGGCATATATCTGATGTGTCCAGCTATTTCTGATTTTATAGACTGTGG	600
Db	541	CTTAATCTCTGGGCATATATCTGATGTGTGTCCAGCTATTTCTGATTTTATAGACTGTGG	600
QY	601	ATCAGTGTCCCTTTTGCCCTGTGCATCTGTGACATCCGAGTACCGAATGGATTCACACTGTG	660
Db	601	ATCAGTGTCCCTTTTGCCCTGTGCATCTGTGACATCCGAGTACCGAATGGATTCACACTGTG	660
QY	661	CATGCTAAATACAGAGTCCCTGCGCTGGGAACATTGAATCAGTTGAAGTCAACCTGG	720
Db	661	CATGCTAAATACAGAGTCCCTGCGCTGGGAACATTGAATCAGTTGAAGTCAACCTGG	720
QY	721	CTGATAAATTTCTGTGTTATGTAGTGTGGGTGGAATCCATGGCAAGCCTACTTCCAGAG	780
Db	721	CTGATAAATTTCTGTGTTATGTAGTGTGGGTGGAATCCATGGCAAGCCTACTTCCAGAG	780
QY	781	GTCCTCTTATCTTACGCCACCTATGCTCAGGTAATGCTCTTCTGTGGCAGCTTTGGG	840
Db	781	GTCCTCTTATCTTACGCCACCTATGCTCAGGTAATGCTCTTCTGTGGCAGCTTTGGG	840

Db	781	GTGCTCTCTTCATCCTCAGCCACCTAATGCTCAGAGTACTGCTCTCTCCAGACTTTTG	840
Qy	841	TGCTGTGTATGGCTCTACACCGGCATATATGATATAGAGCTATTTGAGCTTCCACAGATGG	900
Db	841	TGCTGTGTATGGCTCTACACCGGCATATATGATATAGAGCTATTTGAGCTTCCACAGATGG	900
Qy	901	AACCAAGATCCTCCTAAGGGATATCCAGATCCCAAGACTAAGAGGAGCAGATGATTTCTC	960
Db	901	AACCAAGATCCTCCTAAGGGATATCCAGATCCCAAGACTAAGAGGAGCAGATGATTTCTC	960
Qy	961	CCGATCGTTCTGCAGTACCTCTGCGCTGTGTACATCTCTTTGGGCTTGGTGTGT	1020
Db	961	CCGATCGTTCTGCAGTACCTCTGCGCTGTGTACATCTCTTTGGGCTTGGTGTGT	1020
Qy	1021	TCAGCTCTGTCAATGTCTCAGCTGACTGCTGTCCATCTGTGCGCGAGTTCTATGTGTCT	1080
Db	1021	TCAGCTCTGTCAATGTCTCAGCTGACTGCTGTCCATCTGTGCGCGAGTTCTATGTGTCT	1080
Qy	1081	CGGAAATATCTAACCAAGCTTCTCTTCAGACAAATATGATCAGACAGAAATTTGTGGGTC	1140
Db	1081	CGGAAATATCTAACCAAGCTTCTCTTCAGACAAATATGATCAGACAGAAATTTGTGGGTC	1140
Qy	1141	ATGAGGATCACTGTGCTGTGTGTGGAGCATCTGAGCAACAGCATGGCTTGTGAGAGAG	1200
Db	1141	ATGAGGATCACTGTGCTGTGTGTGGAGCATCTGAGCAACAGCATGGCTTGTGAGAGAG	1200
Qy	1201	ACTGTGTATGGGCTCTGTGTACTGTAGCTCTGACCTTGTCTAATCATCATCTTCCACAG	1260
Db	1201	ACTGTGTATGGGCTCTGTGTACTGTAGCTCTGACCTTGTCTAATCATCATCTTCCACAG	1260
Qy	1261	CTGCTCTGTGTACTCTTTCATCAAAAGAAACCAACATATGGGGCAGTTGCTGTATATTT	1320
Db	1261	CTGCTCTGTGTACTCTTTCATCAAAAGAAACCAACATATGGGGCAGTTGCTGTATATTT	1320
Qy	1321	TTTGGACTATTTCCGAGAAATTACTGAGAGAGCAGCATCTATATCTTGACAGCCCTTATTC	1380
Db	1321	TTTGGACTATTTCCGAGAAATTACTGAGAGAGCAGCATCTATATCTTGACAGCCCTTATTC	1380
Qy	1381	TTTCTACCTGGTATTTACTCTGACAGAAATGTATATACATATGAGGTTCCCATTTTAA	1440
Db	1381	TTTCTACCTGGTATTTACTCTGACAGAAATGTATATACATATGAGGTTCCCATTTTAA	1440
Qy	1441	ACTCTCTCAATGGTACTCTCATTTCTTACCAACATTTGTGTCTTATCTTACGCAAGATAT	1500
Db	1441	ACTCTCTCAATGGTACTCTCATTTCTTACCAACATTTGTGTCTTATCTTACGCAAGATAT	1500
Qy	1501	CTATTTGAAAGTGAAACCTTGCCCTCCAAAAATAGATGATTTTGAATGCTGTGTGCAAG	1560
Db	1501	CTATTTGAAAGTGAAACCTTGCCCTCCAAAAATAGATGATTTTGAATGCTGTGTGCAAG	1560
Qy	1561	CACAGTGAAGAGAACATGAGACAAACATTTCTATGATCGAAATGAAATATTCAAATTTAAAT	1620
Db	1561	CACAGTGAAGAGAACATGAGACAAACATTTCTATGATCGAAATGAAATATTCAAATTTAAAT	1620
Qy	1621	GAACTTGCAACTGTGGAACCTCGGCAABAGCTTAACCTCTAGTTCAACITTCACCAATTAAG	1680
Db	1621	GAACTTGCAACTGTGGAACCTCGGCAABAGCTTAACCTCTAGTTCAACITTCACCAATTAAG	1680
Qy	1681	GAGGCCCTCTTGAATGTGATTTCCAGTCCGAGGGGTCTGGGAGCTGAAGATATTTACAA	1740
Db	1681	GAGGCCCTCTTGAATGTGATTTCCAGTCCGAGGGGTCTGGGAGCTGAAGATATTTACAA	1740
Qy	1741	TGA TGA TGA	1740
Db	1741	TGA TGA TGA	1740
RESULT 2- LOCUS DEFINITION ACCESSION	BC065089 Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds. BC065089	4715 bp mRNA linear ROD 30-JUN-2004	

RESULT 2-				
BC065089				
LOCUS	BC065089	4715 bp	mRNA	linear
DEFINITION	Mus musculus solute carrier family 5 (chooline transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.			
ACCESSION	BC065089			

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

LOCUS	BC065089	4715 bp	mRNA	linear	ROD 30--JUN-2004
DEFINITION	Mus musculus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.				
ACCESSION	BC065089				

VERSION BC065089.1 GI:40787811
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stacieleon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Caerini P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millaby S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Young A.C., Schevchenko Y., Sanchez A., Whiting M., Madan A., Young A.C., Schevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalski U., Smalios D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBLISHED 12477932
REFERENCE 2 (bases 1 to 4715)
AUTHORS Strausberg R.
JOURNAL Direct Submission
TITLE Submitted (02-JAN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: M. Bento Soares, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: Bento-Soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo M.F., Akabogu I., Bair T., Bair J., Crouch K., Davis A., Fishler K., Keppel C., Kucaba T., Lebeck M., Melo A., Schaefer K., Scheetz T., Smith C., Snir E., Tack D., Trout K., Walters J., Casavant T., Soares M.B.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11528509.
FEATURES
source
1. 4715
Location/Qualifiers
1. 4715
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:86023 IMAGE:6851099"
/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
/clone_id="N1H_BMAP_FY0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
1. 4715
/gene="Slc5a7"
/note="SYNonym: CHT1"
/db_xref="GeneID:63993"
/db_xref="MGI:1927126"

CDS
27. 1769
/gene="Slc5a7"
/codon_start=1
/product="solute carrier family 5 (choline transporter), member 7"
/protein_id="AAH65089.1"
/db_xref="GI:40787812"
/db_xref="GeneID:63993"
/db_xref="MGI:1927126"
/translation="MSFVREGIVATILFVLIIFLVIGVIAWAKTKNSGPESEALIV GSRDGLVGFETWATWVGGYINGTBAYVPGCGLAMAPRGYSLSILGSLF AKPMSKQVYMLDPEQOIYGRWGLLFIPLMGEEMMAATFSLALGTSIIIVD VNI5VIVALLAIIYLVGLVSYAVTVVDFICFIIUMISVPSALSHPATVIGFT AVHAKYSPMLGITISVEVYTWLDFLLMGLICFIMQAVFQVLSSTAYQVLSF AAFGLWALPAICGAIAGSDMNOCTAGYVDDPKTEADMLIPIVLOYLCPYISF FGLGVAWSAWSADSITLSSMFRNIVOLSFRONSADKEIVWMTITVPGASA TMAILLITVYGLWATSSDLYIIIFPOLCLVFKGNITGVAVGYIFGLFRLITGG EPYLIQPIIFPGYISDNKGIYNQRFPKLISVTSFTNIVSVLYAKYFESGLP PKLDVDFVAVARHSEENDKTILVRENENIKINELAPVXPROSLTSLSTFNKEALDV DSPSGSTEDNLIQ"
ORIGIN
Query Match 99.4%; Score 1731.8; DB 6; Length 4715;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1736; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
1 ATGCCTTCCATGCTGGAAGGAGCTGGTAGCTATATCTTACCTCTTATTTCTG 60
27 ATGCTTTCCAGTGAAGAGCTGGTAGCTATATCTTACCTCTTATTTCTG 86
61 GTTGAATATGAGGCTGATGTAACCAAAACAGCGGCAACCCAGAGAGCGAGTGA 120
87 GTTGAATATGAGGCTGATGTAACCAAAACAGCGGCAACCCAGAGAGCGAGTGA 146
121 GCATCATATGTCGGGGCCCGTACATGTTGTTGTTGTTGTTTACATGACAGCC 180
147 GCATCATATGTCGGGGCCCGTACATGTTGTTGTTGTTGTTTACATGACAGCC 206
181 ACCGTGGTGGAGAGGCTACATCATGAGGAGCAGAGAGAGAGAGAGAGTGT 240
207 ACCGTGGTGGAGAGGCTACATCATGAGGAGCAGAGAGAGAGAGAGAGTGT 266
241 GGTCTAGCTTGGGCTCATGCAACCATGATGATTTCTGAGTCTAATTTAGTGCTG 300
267 GGTCTAGCTTGGGCTCATGCAACCATGATGATTTCTGAGTCTAATTTAGTGCTG 326
301 TTTTTCGAAACCTATGCTTCCAAAGGATATGATGATGTTAGACCATTTCAACG 360
327 TTTTTCGAAACCTATGCTTCCAAAGGATATGATGATGTTAGACCATTTCAACG 386
361 ATCTATGGAAGGAGAGGAGGCTGCTCATGCTCCGCACTGATGGAGAGATGTC 420
387 ATCTATGGAAGGAGAGGAGGCTGCTCATGCTCCGCACTGATGGAGAGATGTC 446
421 TGGGCTGACGAAATTTCTCTCATTTAGGGGCCACCATGAGCGTATGATGATGAT 480
447 TGGGCTGACGAAATTTCTCTCATTTAGGGGCCACCATGAGCGTATGATGATGAT 506
481 GTGAACATATGCGTCAATGCTCTGCACTCATTTGCCATTTTATACCTAGTGGTGG 540
507 GTGAACATATGCGTCAATGCTCTGCACTCATTTGCCATTTTATACCTAGTGGTGG 566
541 CTCTACTCTGCGGCAATATCTGATGTTGCGAGCTATTTGCAATTTTATAGAGCTGTG 600
567 CTCTACTCTGCGGCAATATCTGATGTTGCGAGCTATTTGCAATTTTATAGAGCTGTG 626
601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGCAATCCGACATCGGATTCACAGCTGTG 660
627 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGCAATCCGACATCGGATTCACAGCTGTG 686
661 CATGCTAATATACAGAGTCCCTGCGGAGACCATTTGATGATGATGATGATGATGAT 720
687 CATGCTAATATACAGAGTCCCTGCGGAGACCATTTGATGATGATGATGATGATGAT 746

QY 721 CTGTATAATTTCTGTATATGATGCTGGTGAATCCATGGAAGCCCTATCCAGAG 780
DB 747 CTTGATATATTTCTGTATATGATGCTGGTGAATCCATGGAAGCCCTATCCAGAG 806
QY 781 GTCCCTCTTATCTCTGAGCCCACTATGCTGAGTACTGCTTCTTGGCAGCTTTTGGG 840
DB 807 GTCCCTCTTATCTCTGAGCCCACTATGCTGAGTACTGCTTCTTGGCAGCTTTTGGG 866
QY 841 TGCCCTGATGAGCTCTACCCGCAATGATGAGAGCTATGAGCTTCCACAGACTGG 900
DB 867 TGCCCTGATGAGCTCTACCCGCAATGATGAGAGCTATGAGCTTCCACAGACTGG 926
QY 901 AACCACTGCTCTACCGGGTATCCAGATCCCAAGACTTAAGAGAAAGCAGATGATTC 960
DB 927 AACCACTGCTCTACCGGGTATCCAGATCCCAAGACTTAAGAGAAAGCAGATGATTC 986
QY 961 CCGATGCTTCTGAGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 987 CCGATGCTTCTGAGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
QY 1021 TCAGCTGCTGATCTCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1047 TCAGCTGCTGATCTCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
QY 1081 CGGAATATCTACGAGCTTCTCTGAGCAAAATGATCATGAGAGAAATGCTGCTGCT 1140
DB 1107 CGGAATATCTACGAGCTTCTCTGAGCAAAATGATCATGAGAGAAATGCTGCTGCT 1166
QY 1141 ATGAGGATCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1167 ATGAGGATCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226
QY 1201 ACTGCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1227 ACTGCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1286
QY 1261 CTGCTCTGCTGATCTCTCTGAGCAAAAGCAACTTATGAGGCTGCTGCTGCTGCT 1320
DB 1287 CTGCTCTGCTGATCTCTCTGAGCAAAAGCAACTTATGAGGCTGCTGCTGCTGCT 1346
QY 1321 TTGAGCATATCTGAGATCTGAGAGAGAGCAATCTATCTGAGGCTGCTGCTGCT 1380
DB 1347 TTGAGCATATCTGAGATCTGAGAGAGAGCAATCTATCTGAGGCTGCTGCTGCT 1406
QY 1381 TTCTACCTGCTGATCTCTGAGCAAAAGCAACTTATGAGGCTGCTGCTGCTGCT 1440
DB 1407 TTCTACCTGCTGATCTCTGAGCAAAAGCAACTTATGAGGCTGCTGCTGCTGCT 1466
QY 1441 ACTCTCTGATGAGTATCTCTGAGCAAAAGCAACTTATGAGGCTGCTGCTGCTGCT 1500
DB 1467 ACTCTCTGATGAGTATCTCTGAGCAAAAGCAACTTATGAGGCTGCTGCTGCTGCT 1526
QY 1501 CTATTTGAAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
DB 1527 CTATTTGAAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1586
QY 1561 CACAGTGAAGAGATGAGCAAAAGCAACTTATGAGGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 1587 CACAGTGAAGAGATGAGCAAAAGCAACTTATGAGGCTGCTGCTGCTGCTGCTGCTGCT 1646
QY 1621 GAACTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 1647 GAACTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1706
QY 1681 GAGGCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 1707 GAGGCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1766
QY 1741 TGA 1743
DB 1767 TGA 1769

RESULT 3
LOCUS BD012720 1743 bp DNA linear PAT 02-AUG-2002
DEFINITION High-affinity choline transporter.
ACCESSION BD012720.1 GI:22092909
VERSION WO 0116315-A/4.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Haga, T. and Okuda, T.
TITLE High-affinity choline transporter
JOURNAL Patent: WO 0116315-A 4 08-MAR-2001;
JAPANESE SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA
COMMENT
OS Mus musculus (mouse)
PN WO 0116315-A/4
PD 08-MAR-2001
PF 18-AUG-2000 WO 2000JP005545
PR 27-AUG-1999 JP 99P 240642, 27-DEC-1999 JP 99P 368991 PI
TATSUYA HAGA, TAKASHI OKUDA
PC C12N5/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC
A61K38/17,
A61K45/00, A61P25/28, G01N33/53, A01K67/027
CC
FH Key Location/Qualifiers
FT CDS (1)..(1743).
FEATURES
source
ORIGIN
Query Match 99.3%; Score 1730.2; DB 2; Length 1743;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1735; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 ATGCTTTCATGCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
DB 1 ATGCTTTCATGCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
QY 61 GTTGAATATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 GTTGAATATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 GCCATCATAGTGGGGGCGCTGACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 121 GCCATCATAGTGGGGGCGCTGACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 121 ACCGCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 121 ACCGCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 GGTCTAGCTGGGCTCATGCAACCATTTGATATCTCTGAGTCTTAATTTAGTGGTCTG 300
DB 241 GGTCTAGCTGGGCTCATGCAACCATTTGATATCTCTGAGTCTTAATTTAGTGGTCTG 300
QY 301 TTTTTCGGAACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 TTTTTCGGAACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 ATCTATGGAAGCGCATGAGTGGGCTGCTCTTCAATCCCTGCACTGATGGAGAGATGCTTC 420
DB 361 ATCTATGGAAGCGCATGAGTGGGCTGCTCTTCAATCCCTGCACTGATGGAGAGATGCTTC 420
QY 421 TGGCTGAGCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 TGGCTGAGCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY	481	GTGAACATATGGGATCTGTCTCGACATCATTTGGCAATTCCTTTATACCCTAGTGGGTGGG	540
Db	481	GTGAACATATGGGATCTGTCTCGACATCATTTGGCAATTCCTTTATACCCTAGTGGGTGGG	540
QY	541	CTTACTCTGTGGCATAATATCATATGTGTCCAGCTATTTCTGCATTTTATATAGACTGTGG	600
Db	541	CTTACTCTGTGGCATAATATCATATGTGTCCAGCTATTTCTGCATTTTATATAGACTGTGG	600
QY	601	ATCAGTGTCCCTTTTGGCCCTGTGCATTCCTTGCAGTCAACCGACATTCGGATTCAAGCTGTG	660
Db	601	ATCAGTGTCCCTTTTGGCCCTGTGCATTCCTTGCAGTCAACCGACATTCGGATTCAAGCTGTG	660
QY	661	CATGCTAAATACAGAGATCCCTGGGTGGGAACATTTGAATCAGTTGAAGTCTAACCTGG	720
Db	661	CATGCTAAATACAGAGATCCCTGGGTGGGAACATTTGAATCAGTTGAAGTCTAACCTGG	720
QY	721	CTTGATTAATTTTCTGTTATATGATGTGGGTGGAAATCCATGGCAAGCTTACTTCCAGAG	780
Db	721	CTTGATTAATTTTCTGTTATATGATGTGGGTGGAAATCCATGGCAAGCTTACTTCCAGAG	780
QY	781	GTGCTCTCTTATCTCTGACGACCTATGCTCAGGAGTATGTCCTTCTGCGACCTTTGGG	840
Db	781	GTGCTCTCTTATCTCTGACGACCTATGCTCAGGAGTATGTCCTTCTGCGACCTTTGGG	840
QY	841	TGCCTGTGATGGCTCTAACCCGCCATATGATAGAGCTATTTGAGCTTCCACAGACTGG	900
Db	841	TGCCTGTGATGGCTCTAACCCGCCATATGATAGAGCTATTTGAGCTTCCACAGACTGG	900
QY	901	AACCGAGCTGGCTTACGGGTATTCAGATATCCCAAGACTAAGAGAGAAAGCAGATATTTCTC	960
Db	901	AACCGAGCTGGCTTACGGGTATTCAGATATCCCAAGACTAAGAGAGAAAGCAGATATTTCTC	960
QY	961	CCGATTCGTTCTGCAGTACTCTGGCCCTGTGTACATCTCTTCTTGGGCTTGGTGTCTGT	1020
Db	961	CCGATTCGTTCTGCAGTACTCTGGCCCTGTGTACATCTCTTCTTGGGCTTGGTGTCTGT	1020
QY	1021	TCAGCTGTGTCAATGTCTCAAGCTGACTGTTCATCCTGTCTGGGAGTTTCTATGTTTGTCT	1080
Db	1021	TCAGCTGTGTCAATGTCTCAAGCTGACTGTTCATCCTGTCTGGGAGTTTCTATGTTTGTCT	1080
QY	1081	CGGAATATCTTACAGAGCTTTCCTTACAGCAAAATGATACAGCAAGAAATGTGGGGTTC	1140
Db	1081	CGGAATATCTTACAGAGCTTTCCTTACAGCAAAATGATACAGCAAGAAATGTGGGGTTC	1140
QY	1141	ATGAGGATCACTGTGCTTGTGTTCGGAGCATCTGGAACAGCCATGAGCTTGTCTGCAGAG	1200
Db	1141	ATGAGGATCACTGTGCTTGTGTTCGGAGCATCTGGAACAGCCATGAGCTTGTCTGCAGAG	1200
QY	1201	ACTGTGTATATGGGCTCTGGTACCTGAGCTGTGACCTTGTCTACATCATATTTTCCACAG	1260
Db	1201	ACTGTGTATATGGGCTCTGGTACCTGAGCTGTGACCTTGTCTACATCATATTTTCCACAG	1260
QY	1261	CTGCTCTGTGACTCTTTCATCAAAAGAAACAACCTTATGGGGGAGTGTGCTGTATATTT	1320
Db	1261	CTGCTCTGTGACTCTTTCATCAAAAGAAACAACCTTATGGGGGAGTGTGCTGTATATTT	1320
QY	1321	TTTGGACTATTTCTTGAGATATTACTGAGAGAGACCATATCTTATACTTGCAGCCCTTATTC	1380
Db	1321	TTTGGACTATTTCTTGAGATATTACTGAGAGAGACCATATCTTATACTTGCAGCCCTTATTC	1380
QY	1381	TTTCAACCCTGGTTTATTACTGTGACAGAAATGATATATCATTCAGAGGTTCCCATTTTAA	1440
Db	1381	TTTCAACCCTGGTTTATTACTGTGACAGAAATGATATATCATTCAGAGGTTCCCATTTTAA	1440
QY	1441	ACTCTCTGCATGTACTCTCAATTCCTTATCCAAATTTGTGTTTCTTATCTAGCCAAAGTAT	1500
Db	1441	ACTCTCTGCATGTACTCTCAATTCCTTATCCAAATTTGTGTTTCTTATCTAGCCAAAGTAT	1500
QY	1501	CTATTTTGAAGTGGAAACCTTGCTCTCCAAAAATTAAGTATTTGAATGCTGTGTGGCAAGG	1560
Db	1501	CTATTTTGAAGTGGAAACCTTGCTCTCCAAAAATTAAGTATTTGAATGCTGTGTGGCAAGG	1560
QY	1561	CACAGTGAAGGAACATGAGACAGACCATCTTCACTCAGAAATGAATAATCAAAATTAAT	1620

[illegible]

```

QY 301 TTTTGGGAAACCTATGCGTTCCAAAGGATATGTACTATGTAGACCATTCACAG 360
Db 301 TTTTGGGAAACCTATGCGTTCCAAAGGATATGTACTATGTAGACCATTCACAG 360
QY 361 ATCTATGAAAGCGATGGGTGGCTCTTCACTCCCGCACTGATGGGAGATGTTTC 420
Db 361 ATCTATGAAAGCGATGGGTGGCTCTTCACTCCCGCACTGATGGGAGATGTTTC 420
QY 421 TGGGTGAGCAATTTCTCTGCACTTAAAGGGCCACCATCAGCGTATGATGTGAT 480
Db 421 TGGGTGAGCAATTTCTCTGCACTTAAAGGGCCACCATCAGCGTATGATGTGAT 480
QY 481 GTGAACATATGCGTATGTTCTCTGCACTCATTTGCCATTTCTTAAACCTAGTGAGT 540
Db 481 GTGAACATATGCGTATGTTCTCTGCACTCATTTGCCATTTCTTAAACCTAGTGAGT 540
QY 541 CTCTACTGTTGGGCAATATGATGTTGTCAGCTATTCCTGATTTTATAGACCTGG 600
Db 541 CTCTACTGTTGGGCAATATGATGTTGTCAGCTATTCCTGATTTTATAGACCTGG 600
QY 601 ATCACTGTCCTTTTGCCCTGTCACATCTGTCAGTCACCGACATGCGATTACAGCTGTG 660
Db 601 ATCACTGTCCTTTTGCCCTGTCACATCTGTCAGTCACCGACATGCGATTACAGCTGTG 660
QY 661 CATGCTAATATACAGAGTCCCTGGCTGGGACCATTTGAATCATGTTGAAGTTACACTGG 720
Db 661 CATGCTAATATACAGAGTCCCTGGCTGGGACCATTTGAATCATGTTGAAGTTACACTGG 720
QY 721 CTGTAATATTTCTGTATATGATGCTGGGTGGAATCCCATGAGCAAGCTACTTCCAGAG 780
Db 721 CTGTAATATTTCTGTATATGATGCTGGGTGGAATCCCATGAGCAAGCTACTTCCAGAG 780
QY 781 GTCTCTCTTCACTCTCAGCCACTATGCTCAGGATCTCTTCTGCGACCTTTGGG 840
Db 781 GTCTCTCTTCACTCTCAGCCACTATGCTCAGGATCTCTTCTGCGACCTTTGGG 840
QY 841 TGCCGTGATGAGCTCTCAACCCCAATGCAATGAGAGCTATTGAGCTTCCACAGACTGG 900
Db 841 TGCCGTGATGAGCTCTCAACCCCAATGCAATGAGAGCTATTGAGCTTCCACAGACTGG 900
QY 901 AACGAGATGCGCTACGGGATCCAGATCCCAAGACTAAGGAGAGACAGATGATTC 960
Db 901 AACGAGATGCGCTACGGGATCCAGATCCCAAGACTAAGGAGAGACAGATGATTC 960
QY 961 CGATCGTTCTGCACTCTGCGCTGTGATCATCTCTTCTTTGGGCTGTGCTGT 1020
Db 961 CGATCGTTCTGCACTCTGCGCTGTGATCATCTCTTCTTTGGGCTGTGCTGT 1020
QY 1021 TCAGTGTCTGATCTCTCACTGCTGATCTGTGCTGATCTGTGCGGAGTTCTAATGTTGCT 1080
Db 1021 TCAGTGTCTGATCTCTCACTGCTGATCTGTGCTGATCTGTGCGGAGTTCTAATGTTGCT 1080
QY 1081 CGGAATATCTACCACTTCTCTCAGACAAATATGATCAGACAAAGAAATGTGTGGT 1140
Db 1081 CGGAATATCTACCACTTCTCTCAGACAAATATGATCAGACAAAGAAATGTGTGGT 1140
QY 1141 ATGAGGATCACTGTCTGTGTTGCGAGATCTGCAACAGCATGCTTGTGTCAGAG 1200
Db 1141 ATGAGGATCACTGTCTGTGTTGCGAGATCTGCAACAGCATGCTTGTGTCAGAG 1200
QY 1201 ACTGTGATGGGCTCTGATCTGAGCTGATCTGTGCTATCATCATCTTCCACAG 1260
Db 1201 ACTGTGATGGGCTCTGATCTGAGCTGATCTGTGCTATCATCATCTTCCACAG 1260
QY 1261 CTGCTCTGTGATCTTCTATCAAGAAACCAACATTTAGGGGCGAGTTCTGTATAT 1320
Db 1261 CTGCTCTGTGATCTTCTATCAAGAAACCAACATTTAGGGGCGAGTTCTGTATAT 1320
QY 1321 TTTGACATATCTCTGAGATTACTGAGAGAGCCATATCTATCTTGAGCCCTTAATC 1380
Db 1321 TTTGACATATCTCTGAGATTACTGAGAGAGCCATATCTATCTTGAGCCCTTAATC 1380

```

```

QY 1381 TTCTACCTGTGTTATTTACTCTGACAGAGATGATATATCAATCAGAGTTCCATTAA 1440
Db 1381 TTCTACCTGTGTTATTTACTCTGACAGAGATGATATATCAATCAGAGTTCCATTAA 1440
QY 1441 ACTCTCTCAGTGTACCTCACTTCTTTACCAACATTTGTGTTCTTATCTAGCCAGAT 1500
Db 1441 ACTCTCTCAGTGTACCTCACTTCTTTACCAACATTTGTGTTCTTATCTAGCCAGAT 1500
QY 1501 CTATTTGAAGTGAACCTTCCCTCCAAAATTAATGATTTGATGCTGTGTGCGAAG 1560
Db 1501 CTATTTGAAGTGAACCTTCCCTCCAAAATTAATGATTTGATGCTGTGTGCGAAG 1560
QY 1561 CACAGTGAAGAGACATGAGAACCACTTTCTAGTCAGAAATGAATAATCAATTAAT 1620
Db 1561 CACAGTGAAGAGACATGAGAACCACTTTCTAGTCAGAAATGAATAATCAATTAAT 1620
QY 1621 GAACCTTGACCTGTGAACCTCGGAGAGCCTAACCCCACTTCAACTTACCAATAG 1680
Db 1621 GAACCTTGACCTGTGAACCTCGGAGAGCCTAACCCCACTTCAACTTACCAATAG 1680
QY 1681 GAGGCCCTCTGTGATGTTGATTCAGTCGAGAGGGGTCTGGGACTGAATATTTACA 1740
Db 1681 GAGGCCCTCTGTGATGTTGATTCAGTCGAGAGGGGTCTGGGACTGAATATTTACA 1740
QY 1741 TGA 1743
Db 1741 TGA 1743

```

RESULT 5
AX080443 4938 bp DNA linear PAT 22-FEB-2001
LOCUS AX080443
DEFINITION Sequence 1 from Patent WO0078950.
ACCESSION AX080443
VERSION AX080443.1 GI:13159872
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 Sierzege, M. and Albrandt, K.
AUTHORS Differentially expressed genes in the adipocytes of obese mice
TITLE Patent: WO 0078950-A 1 28-DEC-2000;
JOURNAL AMYLIN PHARMACEUTICALS, INC. (US)
FEATURES
source location/Qualifiers
1..4938
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
/note="P4P6B1"

ORIGIN

Query Match 99.3%; Score 1730.2; DB 2; Length 4938;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1735; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 1 ATGCTTTCCATGTTGAAGACCTGTAGCTATTTATCTCTTCACTCTTATATTTCTG 60
Db 247 ATGCTTTCCACGTGAAGACCTGTAGCTATTTATCTCTTCACTCTTATATTTCTG 306
QY 61 GTTGAATATGCGCTGCAATGAAAAACAAAAACAGCGCAACCCAGAAAGCGCAGTGA 120
Db 307 GTTGAATATGCGCTGCAATGAAAAACAAAAACAGCGCAACCCAGAAAGCGCAGTGA 366
QY 121 GCAATCATAGTCGGGGGCGGAGCATTTGCTTTGTTGTTGTTGTTTACATGACACCC 180
Db 367 GCAATCATAGTCGGGGGCGGAGCATTTGCTTTGTTGTTGTTTACATGACACCC 426
QY 181 ACTTGGTTGAGAGAGCTACATCAATGGAACAGCAGAAAGAGTATAGGGCAGAGTTGT 240
Db 427 ACTTGGTTGAGAGAGCTACATCAATGGAACAGCAGAAAGAGTATAGGGCAGAGTTGT 486

```

QY 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATCTCTGAGTCTAATTTAGTGTCTG 300
 Db 487 GGTCTAGCTTGGGCTCAGGCAACCATTTGATATCTCTGAGTCTAATTTAGTGTCTG 546
 QY 301 TTTTTCGCAAACTTATGCTTCAAGGATATGATCTATGTTAAGCCATTCAACAG 360
 Db 547 TTTTTCGCAAACTTATGCTTCAAGGATATGATCTATGTTAAGCCATTCAACAG 606
 QY 361 ATCTATGAAAGGCGATGGGCTGCTCTTATCCCTGCACTGATGGAGAGATGTC 420
 Db 607 ATCTATGAAAGGCGATGGGCTGCTCTTATCCCTGCACTGATGGAGAGATGTC 666
 QY 421 TGGGCTGCAACAATTTCTCTGATTAAGGCGCACCATGAGCTGATTCATTTAGTGTGAT 480
 Db 667 TGGGCTGCAACAATTTCTCTGATTAAGGCGCACCATGAGCTGATTCATTTAGTGTGAT 726
 QY 481 GTGAACATATCGGCTATTTGCTCTGCACTTCATTTGCCATTTCTTATTAACCTAGTGGTGG 540
 Db 727 GTGAACATATCGGCTATTTGCTCTGCACTTCATTTGCCATTTCTTATTAACCTAGTGGTGG 786
 QY 541 CTCTACTCTGTCGATATCTGATGTTGTCACGCTATTTCTGCAATTTTATAGACTGTG 600
 Db 787 CTCTACTCTGTCGATATCTGATGTTGTCACGCTATTTCTGCAATTTTATAGACTGTG 846
 QY 601 ATCAGTGTCCCTTTTGGCTGTCACATCCCTGACAGTACCGACATCGGATTCACAGCTGTG 660
 Db 847 ATCAGTGTCCCTTTTGGCTGTCACATCCCTGACAGTACCGACATCGGATTCACAGCTGTG 906
 QY 661 CATGCTAAATACCAGAGTCCCTGCTGGGAAACCATTTGATCAAGTTGAGTCTACACTGTG 720
 Db 907 CATGCTAAATACCAGAGTCCCTGCTGGGAAACCATTTGATCAAGTTGAGTCTACACTGTG 966
 QY 721 CTGTGATATTTCTGTTATTTGATGCTGGGGAATCCAGGAAAGCTATCCAGAG 780
 Db 967 CTGTGATATTTCTGTTATTTGATGCTGGGGAATCCAGGAAAGCTATCCAGAG 1026
 QY 781 GTCTCTCTTTCATCTCAGCAACCTATGCTCAGAGTCTGCTCTTCTGAGCTTTTGG 840
 Db 1027 GTCTCTCTTTCATCTCAGCAACCTATGCTCAGAGTCTGCTCTTCTGAGCTTTTGG 1086
 QY 841 TGCCTGTGATGCTTACCCGCAATATGATGAGTCTATTTGAGCTTTCACAGACTG 900
 Db 1087 TGCCTGTGATGCTTACCCGCAATATGATGAGTCTATTTGAGCTTTCACAGACTG 1146
 QY 901 AACCGAGCTGCTCAGCGGATCCAGATCCCAAGACTAAGAGAGAACAGACATGATTC 960
 Db 1147 AACCGAGCTGCTCAGCGGATCCAGATCCCAAGACTAAGAGAGAACAGACATGATTC 1206
 QY 961 CCGATGCTTCTGAGTACCTGCTGCTGATCATCTCTTCTTGGGCTTGGTGTCTG 1020
 Db 1207 CCGATGCTTCTGAGTACCTGCTGCTGATCATCTCTTCTTGGGCTTGGTGTCTG 1266
 QY 1021 TGAAGCTGTGATGCTCTGAGTGACTGCTGCTCATCTGTGGGAGTTCTATGTTGCT 1080
 Db 1267 TGAAGCTGTGATGCTCTGAGTGACTGCTGCTCATCTGTGGGAGTTCTATGTTGCT 1326
 QY 1081 CGGAATATCTACAGCTTCTTCAAGAAATGCAAGAAAGAAATGATGAGGTC 1140
 Db 1327 CGGAATATCTACAGCTTCTTCAAGAAATGCAAGAAAGAAATGATGAGGTC 1386
 QY 1141 ATGAGAGTACATGCTGTGTTGTTGAGAGATCTGCAACAGCAATGCTTGTGCAAG 1200
 Db 1387 ATGAGAGTACATGCTGTGTTGTTGAGAGATCTGCAACAGCAATGCTTGTGCAAG 1446
 QY 1201 ACTGTGTATGGGCTTGTGATACCTGAGCTTGTCTACATCATCTTCCACAG 1260
 Db 1447 ACTGTGTATGGGCTTGTGATACCTGAGCTTGTCTACATCATCTTCCACAG 1506
 QY 1261 CTGCTGTGTACTCTTCAACAAGAAACCAACATTTAGGGGAGTGTGTTATTT 1320
 Db 1507 CTGCTGTGTACTCTTCAACAAGAAACCAACATTTAGGGGAGTGTGTTATTT 1566

QY 1321 TTGGACTATTCCTGAGAAATTAAGAGAGAGACCATATCTATACCTTGAGCCCTTAATC 1380
 Db 1567 TTGGACTATTCCTGAGAAATTAAGAGAGAGACCATATCTATACCTTGAGCCCTTAATC 1526
 QY 1381 TTCTACCTGCTTATTAATCTGACAAAGATGATATATCAATAGAGGTTCCCATTTAA 1440
 Db 1627 TTCTACCTGCTTATTAATCTGACAAAGATGATATATCAATAGAGGTTCCCATTTAA 1586
 QY 1441 ACTCTCCATGATGATCTCATCTTATTAACAATTTGATGTTCTTATAGCAAGTAT 1500
 Db 1687 ACTCTCCATGATGATCTCATCTTATTAACAATTTGATGTTCTTATAGCAAGTAT 1746
 QY 1501 CTATTTGAAAGTGAACCTTGCTCCAAAATTAAGATGATTTGATGCTGTGCAAG 1560
 Db 1747 CTATTTGAAAGTGAACCTTGCTCCAAAATTAAGATGATTTGATGCTGTGCAAG 1806
 QY 1561 CACAGTGAAGAGAACCTGACAAAGACATTTCTAGTGAAGAAATGAATATCAATTAAT 1620
 Db 1807 CACAGTGAAGAGAACCTGACAAAGACATTTCTAGTGAAGAAATGAATATCAATTAAT 1866
 QY 1621 GAACCTGCACTGTGAACCTGGGAGAGGCTTAACCTCAGTTCACTTACCAATTAAG 1680
 Db 1867 GAACCTGCACTGTGAACCTGGGAGAGGCTTAACCTCAGTTCACTTACCAATTAAG 1926
 QY 1681 GAGGCCCTCTGATGATTTCAAGTCCGAGGGGCTGAGACTGAAGATTAATTAACA 1740
 Db 1927 GAGGCCCTCTGATGATTTCAAGTCCGAGGGGCTGAGACTGAAGATTAATTAACA 1986
 QY 1741 TGA 1743
 Db 1987 TGA 1989

RESULT 5
 MMU401467 1743 bp mRNA linear ROD 16-AUG-2000
 LOCUS
 DEFINITION
 ACCESSION
 AJ401467
 VERSION
 AJ401467.1 GI:9843808
 KEYWORDS
 CHT1 gene; high affinity choline transporter.
 SOURCE
 Mus musculus
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
 1
 Wieland, A., Bonisch, H. and Bruns, M.
 Molecular cloning of the human and murine high affinity choline transporters and characterization of the human gene-structure
 Unpublished
 2 (bases 1 to 1743)
 Bruns, M.
 Direct Submission
 Submitted (14-AUG-2000) Bruns M., University of Bonn, Pharmacology and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY
 Location/Qualifiers
 1..1743
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/cJ"
 /db_xref="taxon:10090"
 /tissue_type="brainstem"
 1..1743
 /gene="CHT1"
 1..1743
 /gene="CHT1"
 /function="sodium- and chloride-dependent reuptake of choline"
 /experiment="experimental evidence, no additional details recorded"
 /codon_start=1
 /product="high affinity choline transporter"

ACCESSION BD012718
 VERSION BD012718.1 GI:22092907
 KEYWORDS WO 0116315-A/2.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciuronomachi; Muridea; Muridae; Murinae; Rattus.
 1 (bases 1 to 1743)
 HAGA, T. and OKUDA, T.
 High-affinity choline transporter
 Patent: WO 0116315-A 2 08-MAR-2001;
 JOURNAL JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA
 OS Rattus norvegicus (rat)
 COMMENT PN WO 0116315-A/2
 PD 08-MAR-2001
 PF 18-AUG-2000 WO 2000JP005545
 PR 27-AUG-1999 JP 99P 240642, 27-DEC-1999 JP 99P 368991 PI
 TATSUYA HAGA, TAKASHI OKUDA
 PC C12N15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC
 A61K38/17
 PC A61K45/00, A61P25/28, G01N33/53, A01K67/027
 CC
 FH Key Location/Qualifiers
 FT CDS Location/Qualifiers
 1..1743
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 ORIGIN
 Query Match 92.7%; Score 1616.6; DB 2; Length 1743;
 Best Local Similarity 95.5%; Pred. No. 0;
 Matches 1664; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Db |||||
 541 CTCTACTCTGTGGCATATAGTATGTGTAACAGTATCTGATTTTATAGAGATGTGG 600
 QY |||||
 Db 601 ATGAGTGTCCCTTTGGCCCTGTCACATCCGTGAGTACCGGATGGATTCACAGCTGG 660
 QY |||||
 Db 601 ATGAGTGTCCCTTTGGCCCTGTCACATCCGTGAGTACCGGATGGATTCACAGCTGG 660
 QY |||||
 Db 661 CATTGCTAAATACAGAGTCCCTGGCTGGGAAACATTGAATCAGTTGAAGTCAACCTGG 720
 QY |||||
 Db 661 CATTGCTAAATACAGAGTCCCTGGCTGGGAAACATTGAATCAGTTGAAGTCAACCTGG 720
 QY |||||
 Db 721 CTGTAAATTTTCTGTATTTGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAGG 780
 QY |||||
 Db 721 CTGTAAATTTTCTGTATTTGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAGG 780
 QY |||||
 Db 781 GTTCTCTCTTCACTCTGAGCCACATAGTGTGAGTACGTCTCTCTGGCAGCTTTGGG 840
 QY |||||
 Db 781 GTTCTCTCTTCACTCTGAGCCACATAGTGTGAGTACGTCTCTCTGGCAGCTTTGGG 840
 QY |||||
 Db 841 TSCCTGTGATGAGCTCTACCCGCAATATGATGAGAGCTATTGAGCTTCCACAGACTGG 900
 QY |||||
 Db 841 TSCCTGTGATGAGCTCTACCCGCAATATGATGAGAGCTATTGAGCTTCCACAGACTGG 900
 QY |||||
 Db 841 TSCCTGTGATGAGCTCTACCCGCAATATGATGAGAGCTATTGAGCTTCCACAGACTGG 900
 QY |||||
 Db 901 AACAGACTGCTTACGGGTATCCAGATCCCAAGACTTAAGAGGAAAGCAGATGATTTCTC 960
 QY |||||
 Db 901 AACAGACTGCTTACGGGTATCCAGATCCCAAGACTTAAGAGGAAAGCAGATGATTTCTC 960
 QY |||||
 Db 961 CCGATGTCTTCCAGTATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 QY |||||
 Db 961 CCGATGTCTTCCAGTATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 QY |||||
 Db 1021 TCAGCTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 QY |||||
 Db 1021 TCAGCTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 QY |||||
 Db 1081 CGGAATATCTACAGCTTCTTCTTCAAGCAAAATGATCCAGCAAGAAATGTGTGGTCTC 1140
 QY |||||
 Db 1081 CGGAATATCTACAGCTTCTTCTTCAAGCAAAATGATCCAGCAAGAAATGTGTGGTCTC 1140
 QY |||||
 Db 1141 ATGAGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 QY |||||
 Db 1141 ATGAGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 QY |||||
 Db 1201 ACTGTGTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 QY |||||
 Db 1201 ACTGTGTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 QY |||||
 Db 1261 CTGCTCTGTGTACTCTTCAACAAAGAACCAACTTATGAGGAGCTGCTGCTGCTGCTG 1320
 QY |||||
 Db 1261 CTGCTCTGTGTACTCTTCAACAAAGAACCAACTTATGAGGAGCTGCTGCTGCTGCTG 1320
 QY |||||
 Db 1321 TTTGAGACTTCTCTGAGAAATTACTGAGAGAGCCATATCTATCTTGAAGCCCTTAATC 1380
 QY |||||
 Db 1321 TTTGAGACTTCTCTGAGAAATTACTGAGAGAGCCATATCTATCTTGAAGCCCTTAATC 1380
 QY |||||
 Db 1381 TTTGAGACTTCTCTGAGAAATTACTGAGAGAGCCATATCTATCTTGAAGCCCTTAATC 1440
 QY |||||
 Db 1381 TTTGAGACTTCTCTGAGAAATTACTGAGAGAGCCATATCTATCTTGAAGCCCTTAATC 1440
 QY |||||
 Db 1381 TTTGAGACTTCTCTGAGAAATTACTGAGAGAGCCATATCTATCTTGAAGCCCTTAATC 1440
 QY |||||
 Db 1441 ACTCTCTCCAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 QY |||||
 Db 1441 ACTCTCTCCAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 QY |||||
 Db 1501 CTAATTGAAAGTGAACCTTGGCTCCAAATTAAGATGATTTGAGAGCTGCTGCTGCTGCTG 1560
 QY |||||
 Db 1501 CTAATTGAAAGTGAACCTTGGCTCCAAATTAAGATGATTTGAGAGCTGCTGCTGCTGCTG 1560
 QY |||||
 Db 1561 CACAGTGAAGAGAACATGAGACAAAGCAATCTAGTCAGAAATGAAGAAATTAATTAAT 1620
 QY |||||
 Db 1561 CACAGTGAAGAGAACATGAGACAAAGCAATCTAGTCAGAAATGAAGAAATTAATTAAT 1620
 QY |||||
 Db 1621 GAACTTGAACCTGCTGGAACCTGGGAGAGCTTACCTTCACTTCACTTCACTTCACTTCA 1680
 QY |||||

Db	1621	GAAGCTGACCTGTAAAGCCTCGACAGAGCCTAACCTCAGTTCACTTCAACAATAA	1680	Qy	361	ATCTATGGAAGCGCGATGGGCTGGCTCTTCAATCCCTGACATGATGGAGATGTTTC	420
Qy	1681	GAGGCCCTCTTGTATGATTCAGTCGCGAGGGGCTCGGACGTGAAGATATTACAA	1740	Db	361	ATCTATGGAAGCGCGATGGGCTGGCTCTTCAATCCCTGACATGATGGAGATGTTTC	420
Db	1681	GAGGCTCCTGTGATGATTCCAGTCAGAGGATCTGGGACTGAAGATTAATAA	1740	Qy	421	TGGGCTGAGCAATTTTCTCTGCAATAGAGGGCCACATCAGCGTATCATGATGTGAT	480
Qy	1741	TGA	1743	Db	421	TGGGCTGAGCAATTTTCTCTGCAATAGAGGGCTACATCAGCGTATCATGATGTGAT	480
Db	1741	TGA	1743	Qy	481	GTGAACATATGGGTATGTTCTCTGCACTATTGGCAATTCCTTATACCTTAGTGGTGG	540
RESULT 8							
E49870		1743 bp	DNA	linear	PAT	27-AUG-2002	
LOCUS	E49870						
DEFINITION	High-affinity choline transporter.						
ACCESSION	E49870.1	GI:22554901					
VERSION	JP 2001136976-A/2.						
KEYWORDS	Rattus sp.						
SOURCE	Rattus sp.						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.						
REFERENCE	Haga, T. and Okuda, T. (bases 1 to 1743)						
AUTHORS	High-affinity choline transporter						
TITLE	Patent: JP 2001136976-A 2 22-MAY-2001;						
JOURNAL	SCIENCE & TECH AGENCY						
COMMENT	OS Rattus sp. (rat)						
	PN JP 2001136976-A/2						
	PD 22-MAY-2001						
	PF 27-DEC-1999 JP 1999368991						
	PI TATSUYA HAGA, TAKASHI OKUDA						
	PC C12N15/09, A01K67/027, A61K38/00, C07K14/47, C07K16/18, C07K19/00,						
	PC C12N5/10,						
	PC C12P21/02, C12P21/08, C12Q1/00, C12N15/00, A61K37/02, C12N5/00 CC						
	FH Key Location/Qualifiers						
	FT CDS (1) . (1743).						
FEATURES	Location/Qualifiers						
source	1. 1743						
	/organism="Rattus sp."						
	/mol_type="genomic DNA"						
	/db_xref="taxon:10118"						
ORIGIN							
Query Match	92.7%;	Score 1616.6;	DB 2;	Length 1743;			
Best Local Similarity	95.5%;	Pred. No. 0;					
Matches 1664;	Conservative 0;	Mismatches 79;	Indels 0;	Gaps 0;			
Qy	1	ATGCCCTTCCATGTGGAAGGAGCTGTGATATATCTTCTTACCTCCTTATTTCTG	60	Qy	1021	TCAGCTGTGTCATGTCCTCAAGCTGATCTGCTCACTCTGTGGGAGTTCTATGTTGCT	1080
Db	1	ATGCCCTTCCATGTGGAAGGAGCTGTGATATATCTTCTTACCTCCTTATTTCTG	60	Db	1021	TCGTGTGTCATGTCCTCAAGCTGATCTGCTCACTCTGTGGGAGTTCTATGTTGCT	1080
Qy	61	GTGGAAATATGGGCTGCATGGAACCAAAACAGCGGCAACCCAGAGGCGCAGTGA	120	Qy	1081	CGGAATATCTACCAAGCTTCTCTGAGACAAATGATCAGACAAGAAATGTGGGCTC	1140
Db	61	GTGGAAATATGGGCTGCATGGAACCAAAACAGCGGCAACCCAGAGGCGCAGTGA	120	Db	1081	CGGAATATCTACCAAGCTTCTCTGAGACAAATGATCAGACAAGAAATGTGGGCTC	1140
Qy	121	GCATCATATGTCGGGGGCGGAGACATGTGTTGTGTGGTTTATCCATGACAGCC	180	Qy	1141	ATGAGGATCACTGTGCTTGTGTCGAGCATCTGCAACAGCCATGTGCTGACGAAG	1200
Db	121	GCATCATATGTCGGGGGCGGAGACATGTGTTGTGTGGTTTATCCATGACAGCC	180	Db	1141	ATGAGGATCACTGTGCTTGTGTCGAGCATCTGCAACAGCCATGTGCTGACGAAG	1200
Qy	181	ACCTGGGTTGGAGGAGGCTACATCAATGGAACAGAGAACAGTATAGGGCAGAGTTGT	240	Qy	1201	ACTGTGATGAGGCTCTGTGTAAGCTGTAAGCTTGTCTATCATATATCTTCCCAAG	1260
Db	181	ACCTGGGTTGGAGGAGGCTACATCAATGGAACAGAGAACAGTATAGGGCAGAGTTGT	240	Db	1201	ACTGTGATGAGGCTCTGTGTAAGCTGTAAGCTTGTCTATCATATATCTTCCCAAG	1260
Qy	241	GATCAGCTTGGGCTCATGCAACCATGGATATTTCTGAGTCTAATTTTATAGTGTCTG	300	Qy	1261	CTGCTCTGTGTAAGCTTCTGTAAGGAAACCAACTTATGAGGAGAGTTGCTGTTATTT	1320
Db	241	GATCAGCTTGGGCTCATGCAACCATGGATATTTCTGAGTCTAATTTTATAGTGTCTG	300	Db	1261	CTGCTCTGTGTAAGCTTCTGTAAGGAAACCAACTTATGAGGAGAGTTGCTGTTATTT	1320
Qy	301	TTTTTTGGAAACCTATGCTTCCAAAGGATATGTGACTATGTAGACCCGTTTCAACAG	360	Qy	1321	TTTTGACTATTTCTGAGAAATTAAGGAGAGGAGCATATCTATCTTACACCCCTTAATC	1380
Db	301	TTTTTTGGAAACCTATGCTTCCAAAGGATATGTGACTATGTAGACCCGTTTCAACAG	360	Db	1321	TTTTGACTATTTCTGAGAAATTAAGGAGAGGAGCATATCTATCTTACACCCCTTAATC	1380
Qy	361	TTTTTTGGAAACCTATGCTTCCAAAGGATATGTGACTATGTAGACCCGTTTCAACAG	360	Qy	1381	TTCTAACCTGTGTTATTAATCTGACAAAGATGATATATCAATTCAGAGGTTCCCATTTAA	1440
Db	361	TTTTTTGGAAACCTATGCTTCCAAAGGATATGTGACTATGTAGACCCGTTTCAACAG	360	Db	1381	TTCTAACCTGTGTTATTAATCTGACAAAGATGATATATCAATTCAGAGGTTCCCATTTAA	1440
Qy	1441	ACTCTCTCATGATGTTACCTCATCTTTTACCAAACTTTGTGTTTATCTAGCCAAATAT	1500	Qy	1441	ACTCTCTCATGATGTTACCTCATCTTTTACCAAACTTTGTGTTTATCTAGCCAAATAT	1500

QY		961	CCGATCGCTTCGAGTACCTCTGCGCCCTGTGACATCTCCCTCTTGGGCTTGGTGCCTT	1020
Db		1184	CCGATGTGTTCTACAGTACCTCTGCGCCCTGTGACATTTCTCTTGGGCTTGGTGCCTT	1243
QY		1021	TCAGCTGCTGCATGATCCTCTAGCTGACCTCTGCATCCCTGTCGGGAGTTCATATTTGCT	1080
Db		1244	TCCTGCTGCTGCATGATCCTCTGCGGCTGACCTCATCCATTCATCAGCAAGTTCCATGTTCCT	1303
QY		1081	CGGAATATCTACCAAGCTTTCCTTCAAGCAAAATGCACTAGACAAAGAAATTGTGTGGTTC	1140
Db		1304	CGGAATATCTACCAAGCTTTCCTTCAAGCAAAATGCATCAGCAAGGAATATGTGTGGGCTC	1363
QY		1141	ATGAGGATCACTGAGCTGTGTGTTGGAGCATCTGCGCAAGGCGATGGCTTGGCTAGCGAAG	1200
Db		1364	ATGAGGATCACTGAGCTGTGTGTTGGAGCATCTGCGCAAGGCGATGGCTTGGCTAGCGAAG	1423
QY		1201	ACTGTGTATGAGGCTCTGTGTTACCTGAGCTCTGACCTGTGTCTACATCATCATCTTCCACAG	1260
Db		1424	ACTGTGTATGAGGCTCTGTGTTACCTGAGCTCTGAGCTGTGATCATCATCATCTTCCACAG	1483
QY		1261	CTGCTCTGTGTACTCTTTCATCAAGGAACCAACACTTATGGGGCAGTTCGTGTTATATTT	1320
Db		1484	CTGCTCTGTGTACTCTTTCATCAAGGAACCAACACTTATGGGGCAAGTTCGTGTTATATTT	1543
QY		1321	TTTGGACTATTTCCCTGAGAAATTAATCGSAGAGAGCCATATCTATACTTSCAGCCCTTAATC	1380
Db		1544	TTTGGACTATTTCCCTGAGAAATTAATCGSAGAGAGCCATATCTATACTTSCAGCCCTTAATC	1603
QY		1381	TTTCTACCTCTGGTTATTTACTCTGCAAGAAATGATATATCAATCAAGGTTCCATTTAAA	1440
Db		1604	TTTCTACCTCTGGTTATTTACTCTGCAAGAAATGATATATCAATCAAGGTTCCATTTAAA	1663
QY		1441	ACTCTCTCCAGTGTACTCTCATCTCTTTCACCAACTTTGTGTGTTCTTATCTHAGCCAAAGAT	1500
Db		1664	ACTCTCTCCAGTGTACTCTCATCTCTTTCACCAACTTTGTGTGTTCTTATCTHAGCCAAAGAT	1723
QY		1501	CTATTTTGAAATGTGAACCTTGCTGCTCCAAAAATTAGATGATTTTGATNGCTGTGTGCAAGG	1560
Db		1724	CTATTTTGAAATGTGAACCTTGCTGCTCCAAAAATTAGATATATTTTGATNGCTGTGTGCAAGG	1783
QY		1561	CACAGTGAAGAAACATGAGACAAGCCATTTGATCAGAAATGAAAATATCAAAATTAAAT	1620
Db		1784	CACAGTGAAGAAACATGAGACAAGCCATTTGATCAGAAATGAAAATATCAAAATTAAAT	1843
QY		1621	GAACTTGAACCTGTGAAAACCTTCGGCAGAGCCTTAACCTCAAGTTCAACTTTTCACAAATPAG	1680
Db		1844	GAACTTGAACCTGTGAAAACCTTCGGCAGAGCCTTAACCTCAAGTTCAAACTTTTCACAAATPAA	1903
QY		1681	GAGGCCCTCTCTTGATGTGATTTCCAGTCCGAGGGGCTGTGGGACTGAAGATAAATTTCAA	1740
Db		1904	GAGGCCCTCTCTTGATGTGATTTCCAGTCCGAGGGGCTGTGGGACTGAAGATAAATTTCAA	
QY		1741	TGA 1743	
Db		1964	TGA 1966	
RESULT 10				
LOCUS	AR268949		1743 bp	DNA
DEFINITION	Sequence 1 from patent US 6500643.		linear	PAT 10-APR-2003
ACCESSION	AR268949			
VERSION	AR268949.1		GI:29699686	
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 1743)			
AUTHORS	Wu,D.-H., Gu,Y., Millard,W.J. and He,Y.-J.			
TITLE	Human high affinity choline transporter			
JOURNAL	Patent: US 6500643-A 1 31-DEC-2002; University of Florida; Gainesville, FL			

FEATURES	Location/Qualifiers
SOURCE	1.. 1743
	/organism="unknown"
	/mol_type="genomic DNA"
Query Match	79.1%; Score 1378.2; DB 2; Length 1743;
Best Local Similarity	86.9%; Pred. No. 0;
Matches 1515; Conservative	0; Mismatches 228; Indels 0; Gaps 0
ORIGIN	
QY	1 ATGCTTTCCATGTGGAGAGACTGTAGCTATTAATTCCTTCTTACCTCTTATATTTCG 60
DB	1 ATGGCTTCCATGTGGAGAGACTGTATCTATCATCATGCTTCTTACTTCTTAATTTCG 60
QY	61 GTTGGAAATATGGGCTGCATGTGAAAAACAAAAACAGCGGCAACCCAGAAAGCGCAGTGA 120
DB	61 GTTGGAAATATGGGCTGCCTGTGAAACCAAAAACATGTGCAGCGCAGAAAGCCACGAA 120
QY	121 GCCATCATATGTCGGGGGCGCTGACATGTGGTTGTTGGTTGGTTTACATGACGCC 180
DB	121 GCCATCATATGTTGGGGCCGAGATTTGGTTATGTTGTGTGTGATTTACATGACAGCT 180
QY	181 ACCGTGGGTGAGAGAGGCTATCATCATGGGACAGAGAAAGCTGATATGGGCGAGTTGT 240
DB	181 ACCGTGGGTGAGAGGGGTATATCATGTGCACAGCTGAAGCAGTTTATATACAGGTTAT 240
QY	241 GGTCTAGCTTGGGCTCATGCAACCCATTGGATTTCTGTAGTCTAATTTTAAGTGCTTG 300
DB	241 GGCTTAGCTTGGGGCTCAGGACCAATTTGATTTGATTTCTTATGCTGATTTTAAAGTGCGT 300
QY	301 TTTTGTGGAAACCTATATGCGTTCCAAAGGATATATGTACTATATGTAGACCCATTCAACAG 360
DB	301 TTTCTTTGAAAAACCTATATCGTTCAAAGGGGATATGTGACCATGTTAGACCCGTTTCAGCA 360
QY	361 ATCTATGAAAGCGCATGGATGGGCTGCTCTTCATCCCTGCACCTGATGAGAGATGTTTC 420
DB	361 ATCTATGAAAGCGATGGGCGGAACTCTCTGTTTATTCCTGCACCTGATGAGAGAAATGTTTC 420
QY	421 TGGGCTGAGCAATTTCTCTGCATTTAAGGGGCGCACCATCAGCGTGATCAATTGATGTGAT 480
DB	421 TGGGCTGAGCAAAATTTCTCTGCTTTGGGAGGCCACATCAGCGTGATCATCGATGTGAT 480
QY	481 GTGAACATATGGGTCATATGTCCTGACCATCATTTGCCATCTTTATACCTATGAGGGGTGG 540
DB	481 ATGCAATTTCTGTATCATATCTCTGCACCTCATGTGCACACTCTGTACACACTGTGGGAGGG 540
QY	541 CTCTACTCTGCGCATATATCTGATGTTGTCCAGCTATTCGTGATTTTATATAGACCTGTGG 600
DB	541 CTCTATTCCTGTGGCCTTACATGATGTGCTGTGACGTCCTTTTGCAATTTTGTAGGGGCTGTGG 600
QY	601 ATCAATGTCCTTTTGGCCCTGTGCACATCTGTGACATCCGAAATGGGATTCACAGCTTGG 660
DB	601 ATCAACGTCCTCCCTTTGGCAATGTGCACATCTGTGAGTCGAGAAATCGGGTTCACATGCTGTG 660
QY	661 CATGTAAATATACAGAGTCCCTGGGCTGGGAAACATTTGATCATGATGATGAATCTTCAACTGG 720
DB	661 CATGCAATATACAAAGCCCGTGGCTGGGAACTGTGTGACTCATCTGAAGCTTACTCTTGG 720
QY	721 CTGTGTAATTTTCTGTATATGTATGTGGTGGATTCGATTCGACGAACTTACTTCAGAGG 780
DB	721 CTGTGTAATTTTCTGTGTTGATGTCTGGGTGGAATCCCATGCGCAAGCATATCTTTCAGAGG 780
QY	781 GTCCCTCTTTCATCTCTGACGACCTATGCTCAAGTGTCTTCTTCTTGGGAGCTTTTCGGG 840
DB	781 GTTCTCTCTTCTTCTCTCAAGCACTTATGCTCAAGTGTCTTCTTCTTGGGAGCTTTTCGGG 840
QY	841 TGCCCTGTATGAGCTCTTACCCGCGCATATGATAGAGAGCTATTGTGAGCTTTCACAGACTGG 900
DB	841 TGCCCTGTATGAGGCTATCCCAAGCCATATGTCATTTGGGGCCATTGTGAGCTTTCACAGACTGG 900
QY	901 AACCAAGATCGCTTACGGGTATCCAAATCCCAAGACTAAGAGAGAAAGCAACATGATTTCT 960
DB	901 AACCAAGATCGCATATGAGGCTTCCAAATCCCAAGACTAAGAGAGCAACATGATATTTTA 960

Oy	961	CGAATCGTTTCGAGTACCTCTGCCCCCTGTGATCACTCTCTTCTTTGGGCTTGGTCTGT	1020
Db	961	CCAAATGTTCTGCACTATCTCTGCCCCGTGTATATTTCTTTCTTTGGTCTGGTGCAGTT	1020
Oy	1021	TCAGTGTGTCATATGCTCCTCAGCTACCTGTCATCCTCTGTGGGAGATTCTATGTTGCT	1080
Db	1021	TCGTGTGCTGTATATCTCATACAGAGATTTCTTCATCTTGTTCAGCAAGTTTCCATGTTTGA	1080
Oy	1081	CGGAATATCTACACAGCTTTCCTTCAGCAAAATGATCAGACAGAAATTGTGGGCTC	1140
Db	1081	CGGAATATCTACACAGCTTTCCTTCAGCAAAATGCTTCGGAACAAAGAAATGTTGGGTT	1140
Oy	1141	ATGAGGATCATCTGTGCTGTGTTCGGAGCAATCTGACAAGCATAGCTTTTGCTACAGAG	1200
Db	1141	ATGCAAAATCACAGTGTGTGTGTGGAGCATCTGCAAGCATAGGCTTGTCTGACGAAA	1200
Oy	1201	ACTGTGTATGGGCTCTGTGTACCTGAGCTGTGACCTTGTCTACATCATCATCTTCCACAG	1260
Db	1201	ACTGTGTATGGGCTCTGTGTACCTGAGCTGTGACCTTGTTCATCATCTTATCTTCCCCAG	1260
Oy	1261	CTGCTCTGTGTACTCTTTCATCAACAGAAACCAACTATAGGGGCACTTGTCTGTATATTT	1320
Db	1261	CTGCTCTGTGTACTCTTGTATTAAGGAAACCAACCTATAGGGGCGGTGGCAGGTATATTT	1320
Oy	1321	TTTGGACATATCTCCGAGAAATTACTGGAGAGAGCCATATCTATCTTGCAACCCCTTATC	1380
Db	1321	TCGTGACCTTCTCTGAGAAATTACTGGAGGAGCCATATCTGTATCTTTCAGCCCTTGATC	1380
Oy	1381	TTCTTACCTGTGTTTACTCTGTGACAGAAATGGTATATCAATCAGAGGTTCCCATTTTAA	1440
Db	1381	TTCTTACCTGTGCTATTTACCTGTATATATGTTATATATATCAGAAATTTTCCATTTTAA	1440
Oy	1441	ACTCTCTCCATGCTTACCTCATTTCTTTTACCAACATTTGTGTTTCTTATCTAGCCAGTAT	1500
Db	1441	ACACTCTCCATGCTTACATCATTTCTTTTACCAACATTTGCATCTCTCATCTGACCAAGTAT	1500
Oy	1501	CTATTTTGAAAGTGAAACCTGTGCTTCCAAATTAAGTATTTGATGATGCTGTGTGGCAAG	1560
Db	1501	CTATTTTGAAAGTGAAACCTGTGCACTTAAATTAATTAATTAATTAATTAATTAATTAATTA	1560
Oy	1561	CACAGTGAAGAAACATGAGACAAACCATTTCTATGACAGAAATGAAATATCAAAATTTAAT	1620
Db	1561	CACAGTGAAGAAACATGATTAAGCAATTTCTTGTCAAAATTAATTAATTAATTAATTAAT	1620
Oy	1621	GAACTTGACCTGTGAAACCTTGGGAGAGCTTAACCTCACTTCACTTCAACCAATTAAG	1680
Db	1621	GAACTTGACCTGTGAAACCTTGGGAGAGCATTAACCTCACTTCACTTCAACCAATTAAG	1680
Oy	1681	GAGGCCCCCTTGAATGTTTCCAGTCCGAGAGGGCTGTGGAGCTGAAGATTAATTTACA	1740
Db	1681	GAGGCCCCCTTGAATGTTTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG	1740
Oy	1741	TGA 1743	
Db	1741	TGA 1743	
RESULT 11			
LOCUS	BD012719	1743 bp	DNA
DEFINITION	High-affinity choline transporter.		
ACCESSION	BD012719		
VERSION	BD012719.1	GI:22092908	
KEYWORDS	WO 0116315-A/3.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo		
REFERENCE	1 (bases 1 to 1743)		
AUTHORS	Haga, T. and Okuda, T.		
TITLE	High-affinity choline transporter		

JOURNAL

Patent : WO 0116315-A 3 08-MAR-2001;
JAPAN SCIENCE AND TECHNOLOGY CORP.,TATSUYA HAGA, TAKASHI OKUDA

COMMENT :

OS Homo sapiens (human)
PN WO 0116315-A/3
PD 08-MAR-2001
PF 18-AUG-2000 WO 2000JP005545
PR 27-AUG-1999 JP 99P 240642,27-DEC-1999 JP 99P 368991 PI
TATSUYA HAGA, TAKASHI OKUDA
PC C12N15/12,C07K14/47,C12Q1/68,C07K19/00,C07K16/18,C12N5/10, PC
A61K38/17,
PC A61K45/00,A61P25/28,G01N33/53,A01K67/027

FEATURES

SOURCE

CC Location/Qualifiers
FH (1):(1743).
FT CDS

1.1743
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 78.9%; Score 1375; DB 2; Length 1743;
Best Local Similarity 86.8%; Pred. No.0; Mismatches 230; Indels 0; Gaps 0;
Matches 1513; Conservative 0;

QY 1 ATGCGCTTCCATGTGGAAGAGACTGGTAGCTATTAATCCCTCTCACTCCCTTATATTTCTG 60
Db 1 ATGGCTTTCCATGTGGAAGAGACTGATAGCTATCATCGTGTCTCACTCTTAATTTTCTG 60

QY 61 GTTGGAAATGCGCTGCATGSAAAACCAAAACACGGCGAACCCAGAAAGCCGAGTAA 120
Db 61 GTTGGAAATGCGCGCTGCCTGGAGAACCAAAACAGTGGCAGCGAGAACGCGAGAA 120

QY 121 GCCATCATATAGCGGGGGCCGTGACATTTGTTGGTTGGTGGTTTACCATGACAGCC 180
Db 121 GCCATCATATAGTTGGTGGCCAGAAATTTGTTATTTGTTGGTGAATTTACATGACAGCT 180

QY 181 ACGTGGGTGGAGAGGCTACATCAATGGAACAGACAGAGCAGTGTATGGCCAGTTGT 240
Db 181 ACGTGGGTGGAGAGGGTATATCAATGACACAGCTGAAGCAGTTATGTACCAAGTTAT 240

QY 241 GGTCTAGCTTGGGCTCATGACCCCATTTGGAATTTCTCGAATCTAATTTTGGTGTG 300
Db 241 GGTCTAGCTTGGGCTCATGACCCCATTTGGAATTTCTCGAATCTAATTTTGGTGTG 300

QY 241 GGCGTAGCTGGGGCTCAGGACCAATTGGAATTTCTTATCTGATTTGATTTTGGTGGCTG 300
Db 241 TTTTGTGGAAACCTATGCGTTCCAAAGGATATGACTATGTATGAGCCATTCAAACAG 360

QY 301 TTTTGTGGAAACCTATGCGTTCCAAAGGATATGACTATGTATGAGCCATTCAAACAG 360
Db 301 TTTTGTGGAAACCTATGCGTTCCAAAGGATATGACTATGTATGAGCCATTCAGCAA 360

QY 361 ATCTATGGAAGCGCATGAGTGGGCTGCTCTTCAATCCCTGCACGTAGTGGAGATGTT 420
Db 361 ATCTATGGAAGAACCATGAGGCGGAGCTCCTGTTTATTCCTGCACGATGGGAAATGTT 420

QY 421 TGGGCTCGAGAAATTTCTCTGCAATTTGAGGGGCCCAATCAGCGGATCATTTGATGAT 480
Db 421 TGGGCTCGAGAAATTTCTCTGCTTTGGAGCCACCAATCAGCGGATATGATGATGAT 480

QY 481 GTGAACAATATGCGATATGTTCTTGCACATCAATTCGCAATTTCTTTATACCTATGAGTGG 540
Db 481 ATGCACAATTTCTGTCATCATCTTGCACTATTGCACTCTGTACACACTGTTGGAGGG 540

QY 541 CTCTACTCTGTGGCAATATCATGATGTTGTCCAGCTATTTGCACTTTTATAGGACTGTG 600
Db 541 CTCTAATCTGTGGGCTCATCATGATGTGTTAGCTCTTTTGCATTTTGTGAGGGCTGTG 600

QY 601 ATCAGTGTCCCTTTTGGCCCTGTCAATCTGCAATCCGACATCCGATCCGATTCACAGCTGTG 660
Db 601 ATCAGCGTCCCTTTTGCATTTGTCAATCTGCAATCCGACATCCGAGATCCGGGTTACAGCTGTG 660

QY 661 CATCTTAATATCCAGAGTCCCTGGCTGGGAACATTAATACGTTGAAGTCTAACCTGG 720
Db 661 CATCCCAATATCCAAAGCGGTGGCTGGGAACGTGTGATCTATCGAAGTCTACTCTTGG 720

```

OY 721 CTTGATATTTTCTGTATTTGATGCTGGGTGAATCCATGGCAAGCTTCTCCAGAG 780
Db 721 CTTGATATTTTCTGTATTTGATGCTGGGTGAATCCATGGCAAGCTTCTCCAGAG 780
OY 781 GTCCCTCTTCAATCTCCAGCCCTATGCTCAGTACGTCCTTCTCCAGCTTTGGG 840
Db 781 GTCCCTCTTCAATCTCCAGCCCTATGCTCAGTACGTCCTTCTCCAGCTTTGGG 840
OY 841 TGCCGTGTATGCTCTTACCCGCTATGCTATGAGAGCTTATGAGCTTCCACAGACTGG 900
Db 841 TGCCGTGTATGCTCTTACCCGCTATGCTATGAGAGCTTATGAGAGCTTCCACAGACTGG 900
OY 901 AACGAGACTGCTAGGGGATCCAGATCCAGACTTAAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 AACGAGACTGCTAGGGGATCCAGATCCAGACTTAAGAGAGAGAGAGAGAGAGAGAGAG 960
OY 961 CCGATGCTTCTGACAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 961 CCGATGCTTCTGACAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
OY 1021 TCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 TCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
OY 1081 CGGAATATCTACAGCTTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 CGGAATATCTACAGCTTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
OY 1141 ATGAGAGATCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1141 ATGAGAGATCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
OY 1201 ACTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db 1201 ACTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
OY 1261 CTGCTGTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1261 CTGCTGTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
OY 1321 TTTGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
Db 1321 TTTGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
OY 1381 TTTGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
Db 1381 TTTGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
OY 1441 ACTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Db 1441 ACTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
OY 1501 CTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1501 CTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
OY 1561 CACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 CACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
OY 1621 GAACTTGAACCTGTAACCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 GAACTTGAACCTGTAACCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
OY 1681 GAGGCTCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1740
Db 1681 GAGGCTCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1740
OY 1741 TGA 1743
Db 1741 TGA 1743

```

```

RESULT 12
CQ722844 1743 bp DNA linear PAT 03-FEB-2004
LOCUS CQ722844
DEFINITION Sequence 8778 from Patent WO02068579.
ACCESSION CQ722844
VERSION CQ722844.1 GI:42283701
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
AUTHORS Kits, such as nucleic acid arrays, comprising a majority of
TITLE humenexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 8778 06-SEP-2002;
FEATURES
PE Corperation (NY) (US)
source Location/Qualifiers
1..1743
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 78.9%; Score 1375; DB 2; Length 1743;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;
OY 1 ATGCTTTCATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 1 ATGCTTTCATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
OY 61 GTTGAATATGAGGCTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 61 GTTGAATATGAGGCTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
OY 121 GGCATATATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 GGCATATATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
OY 181 ACCTGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACCTGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
OY 241 GGTCTAGCTTGGGCTCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 GGTCTAGCTTGGGCTCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
OY 301 TTTTTCGAGAAACCTATGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 TTTTTCGAGAAACCTATGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 361 ATCTATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 ATCTATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
OY 421 TGGGCTGAGAGAAATTTCTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 421 TGGGCTGAGAGAAATTTCTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
OY 481 GTGAACATATGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGAACATATGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
OY 541 CTCTACTGCTGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 CTCTACTGCTGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
OY 601 ATCAGTGTCCCTTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 ATCAGTGTCCCTTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

```

QY 661 CATGTAATATACAGAGTCCCTGCTGGGAAACCTTGAATCAGTTGAAGTCTACACCTGG 720
DB 661 CATGCAAAATACCAAAAGCCGCTGGGAACTGTGACTCATCTGAAGTCTACTCTTGG 720
QY 721 CTGTAATATTTTCTGTATATGATGCTGGTGGAAATCCCAATGGCAAGCCCTACTTCCAGAG 780
DB 721 CTGTAATATTTTCTGTATATGATGCTGGTGGAAATCCCAATGGCAAGCCCTACTTCCAGAG 780
QY 781 GTCTCTCTTCACTCTGAGCCCACTATGTCAGAGTACTCTCTTCTGGCACTTTGGG 840
DB 781 GTCTCTCTTCACTCTGAGCCCACTATGTCAGAGTACTCTCTTCTGGCACTTTGGG 840
QY 841 TGCCGTGATAGGCTCTACCCGCAATATCAGATAGAGAGTATGAGCTCCACAGACTGG 900
DB 841 TGCCGTGATAGGCTCTACCCGCAATATCAGATAGAGAGTATGAGCTCCACAGACTGG 900
QY 901 AACGAGACTGCTTACGGGTATCCAGATCCCAAGACTTAAGAGAGAGAGAGATGATCTC 960
DB 901 AACGAGACTGCTTACGGGTATCCAGATCCCAAGACTTAAGAGAGAGAGAGATGATCTC 960
QY 961 CGGATGCTTCTGAGTACTCTGCTGCTGTGTCATCTCTTCTTGGGCTTGGTCTGT 1020
DB 961 CGGATGCTTCTGAGTACTCTGCTGCTGTGTCATCTCTTCTTGGGCTTGGTCTGT 1020
QY 1021 TCAGCTGCTGATGCTCTGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 TCAGCTGCTGATGCTCTGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 CGGAATATCTTACAGACTTCTTCAAGCAAAATGCTGCGAAGAAATGCTTGGGCT 1140
DB 1081 CGGAATATCTTACAGACTTCTTCAAGCAAAATGCTGCGAAGAAATGCTTGGGCT 1140
QY 1141 ATGAGATATCAGTGTCTGTGTCGAGATCTGCAAGAGCATGCTTCTGTCAGAGAG 1200
DB 1141 ATGAGATATCAGTGTCTGTGTCGAGATCTGCAAGAGCATGCTTCTGTCAGAGAG 1200
QY 1201 ACTGATATGAGGCTGCTGAGTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 ACTGATATGAGGCTGCTGAGTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 CTGCTCTGCTGATCTTCTTCAAGAAACCAACATTTAGGGGAGTTGCTGTTATAT 1320
DB 1261 CTGCTCTGCTGATCTTCTTCAAGAAACCAACATTTAGGGGAGTTGCTGTTATAT 1320
QY 1321 TTTGAGATATCTTCTGAGATTAAGTGAAGAGAGCATATCTATCTTGAAGCTTATC 1380
DB 1321 TTTGAGATATCTTCTGAGATTAAGTGAAGAGAGCATATCTATCTTGAAGCTTATC 1380
QY 1381 TTCTACCCCTGCTTATTAATCTGCAAGAAATGATATTAATCAAGAGTTCCATTTAA 1440
DB 1381 TTCTACCCCTGCTTATTAATCTGCAAGAAATGATATTAATCAAGAGTTCCATTTAA 1440
QY 1441 ACTCTCTCCATGCTTACCTTCTTCAACCAATTTGCTTCTTATCTTACCAAGAT 1500
DB 1441 ACTCTCTCCATGCTTACCTTCTTCAACCAATTTGCTTCTTATCTTACCAAGAT 1500
QY 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTAAGATGATTTGATGCTGTGCGAAG 1560
DB 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTAAGATGATTTGATGCTGTGCGAAG 1560
QY 1561 CACAGTGAAGAGAAATGAGCAAGAACCATTTAGTACAGAAATGAAGATTCAGATTAAT 1620
DB 1561 CACAGTGAAGAGAAATGAGCAAGAACCATTTAGTACAGAAATGAAGATTCAGATTAAT 1620
QY 1621 GAAGTTCGACTCTGTAAGCCTTGGAGAGCTTACCTTCAAGTTCACTTCAAGATTAAG 1680
DB 1621 GAAGTTCGACTCTGTAAGCCTTGGAGAGCTTACCTTCAAGTTCACTTCAAGATTAAG 1680
QY 1681 GAGGCTCTCTGATGATTTGATTCAGTCCGAGAGGGGCTGGGAGCTGAAGATTAATTTAG 1740
DB 1681 GAGGCTCTCTGATGATTTGATTCAGTCCGAGAGGGGCTGGGAGCTGAAGATTAATTTAG 1740

QY 1741 TGA 1743
DB 1741 TGA 1743

RESULT 13

E49871
LOCUS E49871
DEFINITION High-affinity choline transporter.
ACCESSION E49871
VERSION E49871.1 GI:22554902
KEYWORDS JP 2001136976-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1743)
AUTHORS Haga, T. and Okuda, T.
TITLE High-affinity choline transporter
JOURNAL Patent: JP 2001136976-A 3 22-MAY-2001;
SCIENCE & TECH AGENCY
COMMENT OS Homo sapiens (human)
PN JP 2001136976-A/3
PD 22-MAY-2001
PE 27-DEC-1999 JP 1999368991
PI TATSUYA HAGA, TAKASHI OKUDA
PC C12N15/09, A01K67/027, A61K38/00, C07K14/47, C07K16/18, C07K19/00,
PC C12N5/10,
PC C12N15/09, A01K67/027, A61K38/00, C07K14/47, C07K16/18, C07K19/00,
PC C12P21/02, C12P21/08, C12Q1/00, C12N15/00, A61K37/02, C12N5/00 CC
FH Key location/Qualifiers
FT CDS (1) . (1743).

FEATURES
source location/Qualifiers
1..1743
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 78.9%; Score 1375; DB 2; Length 1743;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

QY 1 ATGCTTTCATATGAGAAAGAGTGGATGATATATCTTCTTACTCTCTTATTTCTG 60
DB 1 ATGCTTTCATATGAGAAAGAGTGGATGATATATCTTCTTACTCTCTTATTTCTG 60
QY 61 GTTGAATATGAGGCTGATGAGAAACCAAAACAGCGCAACCAAGAGAGGAGTGA 120
DB 61 GTTGAATATGAGGCTGATGAGAAACCAAAACAGCGCAACCAAGAGAGGAGTGA 120
QY 121 GGCATCATATGTCGGGGGCGGTGACATTTGTTGTTGTTGTTGTTTACATGACAGCC 180
DB 121 GGCATCATATGTCGGGGGCGGTGACATTTGTTGTTGTTGTTTACATGACAGCC 180
QY 181 ACCTGGTGGAGAGAGGCTACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGTGT 240
DB 181 ACCTGGTGGAGAGAGGCTACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGTGT 240
QY 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTAGTGGTCTG 300
DB 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTAGTGGTCTG 300
QY 301 TTTTTCGAAACCTATGCTTCCAAAGGATATGATGATATTTAGCCATTCAAGC 360
DB 301 TTTTTCGAAACCTATGCTTCCAAAGGATATGATGATATTTAGCCATTCAAGC 360
QY 361 ATCTATGAAAGCGAGTGGGCTGCTCTTCACTGCACTGATGGAGAGATGTTTC 420
DB 361 ATCTATGAAAGCGAGTGGGCTGCTCTTCACTGCACTGATGGAGAGATGTTTC 420
QY 421 TGGGCTGAGCAATTTTCTCTGCAATTAAGGGCAGACATACAGGTATCATTTGAT 480
DB 421 TGGGCTGAGCAATTTTCTCTGCAATTAAGGGCAGACATACAGGTATCATTTGAT 480

Db 421 TGGGTCGACGAATTTCTCTCTTGGGAGCCACATCAGCTATCATGATGAT 480
Qy 481 GTGAACATATCGGTATGTCCTCTGCACTCATTTGCAATCTTTATACCTTATGTTGGG 540
Db 481 ATGACATATTCGTGATCATCTCTGACATCATTTGCACTCTGTACACTGGTGGAGGG 540
Qy 541 CTCTACTCTGGGCAATATACGATGTTTCCAGCTATTTCTGATTTTTATAGACTGG 600
Db 541 CTCTACTCTGGGCAATATACGATGTTTCCAGCTATTTCTGATTTTTATAGACTGG 600
Qy 601 ATCATGTCCTCTTGGCCCTGTCATCATCTGCAATGACGACATGCAATTCACAGCTGG 660
Db 601 ATCATGTCCTCTTGGCCCTGTCATCATCTGCAATGACGACATGCAATTCACAGCTGG 660
Qy 661 CATGCTAAATACAGAGTCCCTGGGAGACCATTTGAATCAGTTGAAGTCTACCTGG 720
Db 661 CATGCTAAATACAGAGTCCCTGGGAGACCATTTGAATCAGTTGAAGTCTACCTGG 720
Qy 721 CTGTATATTTTCTGTGTTATGATGCTGGGTGGAAATCCCATGGCAAGCCCTACTTCAGAGG 780
Db 721 CTGTATATTTTCTGTGTTATGATGCTGGGTGGAAATCCCATGGCAAGCCCTACTTCAGAGG 780
Qy 781 GTCTCTCTGATCCTCTGACGCACTATGCTCAGTACTGCTCTTCTGGCAGCTTTGGG 840
Db 781 GTCTCTCTGATCCTCTGACGCACTATGCTCAGTACTGCTCTTCTGGCAGCTTTGGG 840
Qy 841 TGCTGTGATGCTCTACCCGCAATGATGATAGAGCTATTTGAGCTTCCACAGACTGG 900
Db 841 TGCTGTGATGCTCTACCCGCAATGATGATAGAGCTATTTGAGCTTCCACAGACTGG 900
Qy 901 AACCAAGATGCTGAGGATCCAGATCCCAAGACTATAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 AACCAAGATGCTGAGGATCCAGATCCCAAGACTATAGAGAGAGAGAGAGAGAGAGAG 960
Qy 961 CCGATGCTTCTGACAGTACCTGCTGCTGTATCATCTCTTCTGGGCTTGGCTGCTGT 1020
Db 961 CCGATGCTTCTGACAGTACCTGCTGCTGTATCATCTCTTCTGGGCTTGGCTGCTGT 1020
Qy 1021 TCACTGCTGTGATCTCTGACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 TCACTGCTGTGATCTCTGACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 1081 CGGAATATCTACCACTCTTCTGACGAAATGCAATGCAAGAAAGAAATGTTGGGCT 1140
Db 1081 CGGAATATCTACCACTCTTCTGACGAAATGCAATGCAAGAAAGAAATGTTGGGCT 1140
Qy 1141 ATGAGGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 ATGAGGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy 1201 ACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 ACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Qy 1261 CTGCTGCTGATCTCTTCAATCAAAAGAACCACTTATGAGGAGCTGCTGCTGCTGCTGCT 1320
Db 1261 CTGCTGCTGATCTCTTCAATCAAAAGAACCACTTATGAGGAGCTGCTGCTGCTGCTGCT 1320
Qy 1321 TTTGGACTATTTCTGAGATTAATCTGAGAGAGAGCCATATCTTATCTTGGAGCCCTTAAATC 1380
Db 1321 TTTGGACTATTTCTGAGATTAATCTGAGAGAGAGCCATATCTTATCTTGGAGCCCTTAAATC 1380
Qy 1381 TTTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Db 1381 TTTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Qy 1441 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 1441 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Qy 1501 CTATTGGAAGTGAAGCTTGGCTCAAAATTTAGATTTGATGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 1501 CTATTGGAAGTGAAGCTTGGCTCAAAATTTAGATTTGATGCTGCTGCTGCTGCTGCTGCTGCT 1560

Qy 1561 CACAGTGAAGAAATGAGCAAGACCAATTTAGTACAGAAATGAAATATCAATTAAT 1620
Db 1561 CACAGTGAAGAAATGAGCAAGACCAATTTAGTACAGAAATGAAATATCAATTAAT 1620
Qy 1621 GAACCTGACCTGTGAAACCTGCGAGAGCCTTACCCCTCACTTCACTTCAAGCAATAG 1680
Db 1621 GAACCTGACCTGTGAAACCTGCGAGAGCCTTACCCCTCACTTCACTTCAAGCAATAG 1680
Qy 1681 GAGGCTCTCTGATGTTGATTTCCAGTCGAGAGGGCTCTGAGCTGAGATTAATTTACAA 1740
Db 1681 GAGGCTCTCTGATGTTGATTTCCAGTCGAGAGGGCTCTGAGCTGAGATTAATTTACAG 1740
Qy 1741 TGA 1743
Db 1741 TGA 1743
RESULT 14
AF276871 1743 bp mRNA linear PRI 27-NOV-2000
LOCUS Homo sapiens high affinity choline transporter (SLC5A7) mRNA,
DEFINITION complete cds.
ACCESSION AF276871 GI:10998441
KEYWORDS AF276871.1 GI:10998441
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo
1 (bases 1 to 1743)
REFERENCE Apparsundaram, S., Ferguson, S.M., George, A.L. Jr. and Blakely, R.D.
AUTHORS Molecular cloning of a human, hemicholinium-3-sensitive choline
TITLE transporter
JOURNAL Biochem. Res. Commun. 276 (3), 862-867 (2000)
PUBMED 11027560
REFERENCE 2 (bases 1 to 1743)
AUTHORS Apparsundaram, S., Ferguson, S.M. and Blakely, R.D.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-2000) Department of Pharmacology and Center for
Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at
Pierce, Nashville, TN 37232-6420, USA
FEATURES
source
1..1743
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q12"
1..1743
/gene="SLC5A7"
1..1743
/gene="SLC5A7"
/note="hCHT: solute carrier family 5 member 7"
/codon_start=1
/product="high affinity choline transporter"
/protein_id="AA025940.1"
/db_xref="GI:10998442"
/translation="MAFVBEGLIAIIVFYLLILVGIWAARTKNSGAESEALIV
GGRDILGVGFTMTATWVGGYNGRAAVYVPGGLAMAOAPIGYSLILGLTF
AKPMRSKGYVMDLPFOQIYKRMGGFLFPAALMGEMFWAAATISALGATISVIIDV
WHISVILSALFATITLVGLSVAYTDVOLFCEVGLMISVFPALSHPAVDIGFT
AVHAKYQPMWGLYDSEVYSWDSFILMGGI.PWQAYFQRLVSSSATAVOYLSTL
AALFCLVWALPAILIGAIASSTDMNQTAVGLPDKTTEADMIIPIVQYCPVYISF
FGLGAVSAVMSADSILSASMSFARNIYOLSERONASDEKIVWMIITVEFGASA
TAMLLIKTVYGLMYSLSDIYIVIPQILCVLFVKGNTRYGAAGVYSGI.FLRTGG
EPYILQPLIYFPGYIPDNGIYNQKPFKTLAMVTSFLTNICISYLAKYIPESGTLF
PKLDVFDVAVARSHSEENMDKITLVKENITKIDELALVAPROSMTLSTFTNKEALDV
DSSESGGTEDNLD"

ORIGIN
Query Match

78.9%; Score 1375; DB 5; Length 1743;

Best Local Similarity 86.8%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

QY 1 ATGCTTTCCATGTTGAAAGAGCTGGTATTAATCTTCTTACCTCTTATTTTCTG 60
DB 1 ATGCTTTCCATGTTGAAAGAGCTGATGATATCATCTGTTCTACCTTTATTTTCTG 60
QY 61 GTTGAATATGAGGCTGCAATGAAAAACAAAAACAGGCGAACCCAGAAAGGCGAGTAA 120
DB 61 GTTGAATATGAGGCTGCTGAGAAACAAAAACAGGCGCGACAGAGAGCGACGA 120
QY 121 GCCATCATAGTGGGCGCGTGAATGGTTTGTGGTGTGGTTTACATGACAGCC 180
DB 121 GCCATCATAGTGGTGGCGGAGATATGGTTATTGGTTGGTGGATTACATGACAGCT 180
QY 181 ACCTGGGTTGAGAGGCTACATCAATGGGACAGAGAGAGAGTGTATGGGCAAGTT 240
DB 181 ACCTGGGTTGAGAGGCTATCAATGAGCAGAGCTGAGAGAGTTATGATCAGGTTAT 240
QY 241 GGCTAGCTTGGGCTCATGACCCCATTTGATATCTGAGTCTAAATTTAGGCTG 300
DB 241 GGCTAGCTTGGGCTCATGACCCCATTTGATATCTGAGTCTAAATTTAGGCTG 300
QY 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGACTATGTAGACCCATTCAAAAG 360
DB 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGACTATGTAGACCCATTCAAAAG 360
QY 361 ATCTATGAAAGCGATGGGTGGGCTGCTTCTTCAATCCCTGCACTGATGGGAAGATGTT 420
DB 361 ATCTATGAAAGCGATGGGTGGGCTGCTTCTTCAATCCCTGCACTGATGGGAAGATGTT 420
QY 421 TGGGCGGAGCAATTTCTGATATGAGGCGCACCATGAGGCTGATCTGATGATGAT 480
DB 421 TGGGCGGAGCAATTTCTGCTGTTGGAGCCACATGAGGCTGATCTGATGATGAT 480
QY 481 GTGAACATATCGGTATGCTCTGCACTCATTTGCCATTTCTTAAACCTTAGTGGTGG 540
DB 481 ATGACATATTCGTATCATCTCTGCACTCATTTGCCATTTCTTAAACCTTAGTGGTGG 540
QY 541 CTCTACTCTGTGGCATATCTGATGTGTCAGCTATTTGCAATTTTATAGACCTGTGG 600
DB 541 CTCTACTCTGTGGCATATCTGATGTGTCAGCTATTTGCAATTTTATAGACCTGTGG 600
QY 601 ATCAAGTCCCTTTGGCCCTGATCATCTGCACTGACAGTACAGGATTTCAAGCTGG 660
DB 601 ATCAAGTCCCTTTGGCCCTGATCATCTGCACTGACAGTACAGGATTTCAAGCTGG 660
QY 661 CATGCTAAATACCAAGTCCCTGCTGGAGAACATTTGATCAGTTGAGTCAACTTGG 720
DB 661 CATGCTAAATACCAAGTCCCTGCTGGAGAACATTTGATCAGTTGAGTCAACTTGG 720
QY 721 CTGATATATTTCTGTTATGATGCTGGGTGGAATCCATGCGAAGCTTATCCAGAG 780
DB 721 CTGATATATTTCTGTTATGATGCTGGGTGGAATCCATGCGAAGCTTATCCAGAG 780
QY 781 GTTCCTCTTCTTCTTCTGACGACCTATGCTCAAGTCTTCTTCTGAGAGCTTTGG 840
DB 781 GTTCCTCTTCTTCTTCTGACGACCTATGCTCAAGTCTTCTTCTGAGAGCTTTGG 840
QY 841 TGCCTGTGATGAGCTCTACCCGCAATGATCATGAGAGTATGAGCTTCCACAGACTGG 900
DB 841 TGCCTGTGATGAGCTCTACCCGCAATGATCATGAGAGTATGAGCTTCCACAGACTGG 900
QY 901 AACCGAGCTGCTACGCGGTATCCAGATCCAGATCCAGATCCAGATCCAGATCCAG 960
DB 901 AACCGAGCTGCTACGCGGTATCCAGATCCAGATCCAGATCCAGATCCAGATCCAG 960
QY 961 CCGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 CCGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 TAGCTGCTGATGCTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 TAGCTGCTGATGCTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

DB 1021 TCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 CGGATATCTACACAGCTTCTCTGAGCAAAATGATCAGACAGAAATTTGTTGGTTC 1140
DB 1081 CGGATATCTACACAGCTTCTCTGAGCAAAATGATCAGACAGAAATTTGTTGGTTC 1140
QY 1141 ATGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1141 ATGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 ACTGTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 ACTGTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 CTGCTGCTGATCTCTGATCAAAAGAACCAACATTTAGGCGAGTTCGCTGATAT 1320
DB 1261 CTGCTGCTGATCTCTGATCAAAAGAACCAACATTTAGGCGAGTTCGCTGATAT 1320
QY 1321 TTTGACTATTTCTGAGAAATTAAGTGAAGAGAGCAATCTATCTGACAGCCCTTATC 1380
DB 1321 TCTGGCTCTTCTGAGAAATTAAGTGAAGAGAGCAATCTATCTGATCTGACCCCTGATC 1380
QY 1381 TTTTACCTGCTGATTTACTCTGACAAAGATGATATCAATCAAGAGTTCCATTAA 1440
DB 1381 TTTTACCTGCTGATTTACTCTGACAAAGATGATATCAATCAAGAGTTCCATTAA 1440
QY 1441 ACTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 ACATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 CTATTTGAAAGTGAACCTTCTGCTCCAAATTTAGATTTGATGCTGCTGCTGCAAG 1560
DB 1501 CTATTTGAAAGTGAACCTTCTGCTCCAAATTTAGATTTGATGCTGCTGCTGCAAG 1560
QY 1561 CACAGTGAAGAGACATGAGCAAGACCATTTAGTCAAGATGAAATTTCAATTAAT 1620
DB 1561 CACAGTGAAGAGACATGAGCAAGACCATTTAGTCAAGATGAAATTTCAATTAAT 1620
QY 1621 GAGCTTGAACCTGTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 1621 GAGCTTGAACCTGTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1681 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 TGA 1743
DB 1741 TGA 1743

RESULT 15
HSA401466 1813 bp mRNA linear PRI 16-AUG-2000
LOCUS HSA401466
DEFINITION Homo sapiens mRNA for high affinity choline transporter (CHT1 gene).
ACCESSION AJ401466
VERSION AJ401466.1 GI:9843753
KEYWORDS CHT1 gene; high affinity choline transporter.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Wieland, A., Bonisch, H. and Brues, M.
Molecular cloning of the human and murine high affinity choline transporters and characterization of the human gene-structure
Unpublished
2 (bases 1 to 1813)
Brues, M.
AUTHORS Brues, M.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2000) Brues M., University of Bonn, Pharmacology

FEATURES
source and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY
location/Qualifiers
1. .1813
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q11-13"
/tissue_type="hypothalamus"
1. .1813
/gene="CHT1"
19. .1761
/gene="CHT1"
/function="sodium- and chloride-dependent reuptake of choline"
/experiment="experimental evidence, no additional details recorded"
/product="high affinity choline transporter"
/protein_id="CAC03717.1"
/db_xref="GI:9843754"
/db_xref="GOA:O9GZV3"
/db_xref="UniProtKB/TrEMBL:O9GZV3"
/translation="MARVELGILAIIVYVILLIVIGIWAARTKNSGASRRSEALIV
GARDIGLVGFTMTATVGGSTINGTIEAVYVPEYGLMAOAPIGYLSLIGLFF
AKPMRSKYVTLDPFOIYGRMGGLFIPALMGEMWAAIPFALATLSIIVD
MHSIVISALATLTVGLGVSVATDVVOLFCIPGLMISVPALSHPAVADIGFT
AHAVAYOKPMLGTVDSEVSWLDSFLMGLGIPMOAYFGVLSSSATPAQVLSFL
AAGGLVMAIPALIGAGSDMNOTFLVGLPDPKTEBADMLPIVLYCPVYISF
FGIGVSAVMSADSILSASMFARNIYOLSPRONADKEIWMVMIIVFVGRASA
TMALITKTVYGLWYSDIYIYIPOLLCTLPKGINTYGAAGVYSGLFRLITG
EPTLYLPLIFPYGTPDDNGIYNOKFPKTLAVTSPLINICISLYLKYIPESGTLF
PKLDVFDVAVHSEEMDKTILVKNENIKDELALVXPROSMTLSFTFMEKFLDV
DSPEGSGTEBDNLQ"

ORIGIN

Query Match 78.9%; Score 1375; DB 5; Length 1813;
Best Local Similarity 86.8%; Pred. No. 0;

Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

1 ATGCTTTCATGTGGAAGACCTGGTAGCTATTATCTTCTACCTCCTTATATTTCTG 60
19 ATGCTTTCATGTGGAAGACCTGGTAGCTATTATCTTCTACCTCCTTATATTTCTG 78
61 GTTGAATATGCGCTGATGAAAAACAACCAAGCGGCAACCCAGAAAGCGCAGTGA 120
79 GTTGAATATGCGCTGATGAAAAACAACCAAGCGGCAACCCAGAAAGCGCAGTGA 138
121 GCCATCATATCGGGGCGGTGACATGTTGTTGTTGTTGTTTACATGACAGCC 180
139 GCCATCATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 198
181 ACCTGGTGTGAGAGAGCTACATCATGGAAGACAGAGAGAGAGAGAGAGAGAGAG 240
199 ACCTGGTGTGAGAGAGAGAGATATCATGAGCAGCAGAGAGAGAGAGAGAGAGAG 258
241 GGTCTAGCTGGGCTCATGACCACTGGATATTTCTCTGAGCTTAATTTTACGCTG 300
259 GGCTAGCTGGGCTCATGACCACTGGATATTTCTCTGAGCTTAATTTTACGCTG 318
301 TTTTGTGGAACCTATGCTTCAAGAGATATGACATGTTTGAACCATTCACAAACG 360
319 TTTTGTGGAACCTATGCTTCAAGAGATATGACATGTTTGAACCATTCACAAACG 378
361 ATCTATGGAAGCGAGTGGGCTGCTTCATCCCTGACAGTATGGAGAGAGATGTT 420
379 ATCTATGGAAGCGAGTGGGCGGAGCTCTGTTTATTCCTGCACTGATGGAGAAATGTT 438
421 TGGGCTGAGCAATTTTCTCTGCAATTAAGGCGCAACATCAGCGTATCATTTGATGAT 480
439 TGGGCTGAGCAATTTTCTCTGCTTGGAGGCAACATCAGCGTATCATTTGATGAT 498
481 GTGAACATATGCGTATGTTCTCTGCACTCATTTGCAATTTTATACCTATGAGGCTG 540

499 ATGCAATTTCTGTCATCATCTCTGACATTCGCACTCTGTCACACTGCTGGAGAGG 558
541 CTCTACTCTGAGGATATATGATGTTGTCAGGATATCTGATTTTATAGACTGTG 600
559 CTCTATCTGAGGCTTACATGATGATGTTGATGATGATGATGATGATGATGATGAT 618
601 ATCACTGCTCTTTTGGCTGCTGATCATCTGCACTGCACTGCACTGCACTGCTG 660
619 ATCACTGCTCTTTTGGCTGCTGATCATCTGCACTGCACTGCACTGCACTGCTG 678
661 CATGCTAATACAGAGTCCCTGCTGGAACATGATGATGATGATGATGATGATGAT 720
679 CATGCTAATACAGAGTCCCTGCTGGAACATGATGATGATGATGATGATGATGAT 738
721 CTGCTAATTTTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 780
739 CTGCTAATTTTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 798
781 GTCTCTCTCATCTCTGAGCACTGATGATGATGATGATGATGATGATGATGATGAT 840
799 GTCTCTCTCTCTCTCTGAGCACTGATGATGATGATGATGATGATGATGATGATGAT 858
841 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
859 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 918
901 AACCACTGCTGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
919 AACCACTGATATATGAGCTTCAAGATCAGAACTAGAAAGAGGAGAGATGATTTTA 978
961 CGGATGCTTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
979 CGGATGCTTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
1021 TCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1039 TCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
1081 CGGATATCTACAGCTTCTCTGAGCACTGATGATGATGATGATGATGATGATGATGAT 1140
1099 CGGATATCTACAGCTTCTCTGAGCACTGATGATGATGATGATGATGATGATGATGAT 1158
1141 ATGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
1159 ATGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218
1201 ACTGATATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
1219 ACTGATATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1278
1261 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
1279 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1338
1321 TTTGAGATATCTGAGAAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
1339 TCTGAGCTCTCTCTGAGAAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1398
1381 TTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
1399 TTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1458
1441 ACTCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
1459 ACCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1518
1501 CTATTTGAAAGTGAACCTTGGCTTCAAAATTAATGATGATGATGATGATGATGATG 1560
1519 CTATTTGAAAGTGAACCTTGGCTTCAAAATTAATGATGATGATGATGATGATGATG 1578
1561 CACAGTGAAGAGATGAG 1620

Db 1579 CACAGTGAAGAAACATGATAGACAATTCTTGCAAAATGAAATATTAATTAGAT 1638
QY 1621 GAACTTGCACTGTGAAACCTCGGACAGAGCTTAACCTCAGTTCAACTTTCAACAATAG 1680
Db 1639 GAACTTGCACTGTGAGCCACGACAGACATGACCTCAGCTCAACTTTCAACAATATA 1698
QY 1681 GAGGCCCTCTGATGTTGATTCAGTCCGAGAGGGCTGAGACTGAAGATAATTACAA 1740
Db 1699 GAGGCCCTCTGATGTTGATTCAGTCCAGAAAGGCTGAGACTGAAGATAATTACAG 1758
QY 1741 TGA 1743
Db 1759 TGA 1761

Search completed: July 11, 2006, 02:51:36
Job time : 9697 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 10, 2006, 14:03:03 ; Search time 9507 Seconds
(without alignments)
5851.926 Million cell updates/sec

Title: US-10-724-806-4
Perfect score: 2993
Sequence: 1 MPFHVEGVAILFLYLLFL.....EALLDVDSPEGSGETDNLQ 580

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 segs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seg length: 0
Maximum DB seg length: 200000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus_p2n.model -DEV=xlh
-Q=/abs/ABSWEB.spool/US10724806/runat.10072006.140243.7656/app_query.fasta.1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pep -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pep -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03h
-USER=US10724806@CCN_1.1.5194@runat.10072006.140243.7656 -NCRP=6 -ICPU=3
-NO_MAP -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV.TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*

1:	gb_env:*
2:	gb_pat:*
3:	gb_ph:*
4:	gb_pl:*
5:	gb_pri:*
6:	gb_ro:*
7:	gb_sts:*
8:	gb_sy:*
9:	gb_un:*
10:	gb_vl:*
11:	gb_ov:*
12:	gb_hcg:*
13:	gb_in:*
14:	gb_om:*
15:	gb_da:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2993	100.0	1743	6	AF276872 Mus muscu
2	2973	99.3	4715	6	BC065089 Mus muscu
3	2968	99.2	1743	2	BD012720 Hsgh-affi

Result No.	Score	Query Match	Length	DB ID	Description
4	2968	99.2	1743	2	E49872 High-affini
5	2967	99.1	4938	2	AX080443 Sequence
6	2958	98.8	1743	6	MM4041467 Mus muscu
7	2956	98.8	1743	2	BD012718 High-affi
8	2956	98.8	1743	2	E49870 High-affi
9	2956	98.8	4904	6	AB030947 Rattus no
10	2795	93.4	1743	2	BD012719 High-affi
11	2795	93.4	1743	2	C0722844 Sequence
12	2795	93.4	1743	2	E49871 High-affi
13	2795	93.4	1743	2	AR268949 Sequence
14	2795	93.4	1743	5	AF276871 Homo sapi
15	2795	93.4	1813	5	HS4041466 Homo sapi
16	2795	93.4	2583	5	BC111525 Homo sapi
17	2795	93.4	5158	5	AB043997 Homo sapi
18	2695	90.0	2541	5	BC111524 Homo sapi
19	2680	89.5	2112	4	BC112794 Bos tauru
20	2226	74.4	2528	11	TM4420808 Torpede m
21	1696	56.7	16600	11	AC174718 Gasterost
22	1694	56.6	1132	11	GG4511267 Gallus ga
23	1542.5	51.5	3255	13	AY047521 Drosophi
24	1517	50.7	3326	13	AY011119 Limulus p
25	1489.5	49.8	2049	13	AY629593 Trichopl
26	1444.5	48.3	1731	2	BD012717 High-affi
27	1444.5	48.3	1731	2	E49869 High-affi
28	1444.5	47.2	1985	13	AB030946 Caenorhab
29	1413	47.2	4223	13	CO613835 Sequence
30	1413	47.2	186989	13	AC007812 Drosophi
31	1413	47.2	189116	13	AC009395 Drosophi
32	1413	47.2	255619	13	AE003723 Drosophi
33	1405	46.9	1729	2	CO613836 Sequence
34	1386	46.3	140156	12	AC017381 Drosophi
35	1200.5	40.1	155210	6	AC102873 Mus muscu
36	1200.5	40.1	257456	12	CT571242 Mus muscu
37	1152.5	38.5	40893	13	CBRG45E19 Caenorhab
38	1136	38.0	232792	12	AC106657 Rattus no
39	1105.5	36.9	33908	13	CRC48D1 Caenorhabd
40	1103	36.9	110000	15	AE017333_31 Continuation (32 o
41	1103	36.9	110000	15	CP000002_31 Continuation (32 o
42	1099.5	36.7	1461	2	AX432086 Sequence
43	1097	36.7	159918	12	AC182392 Pan trogl
44	1096	36.6	190043	5	AC009963 Homo sapi
45	1094.5	36.6	330724	12	CEY67H2 Caenorhab

ALIGNMENTS

RESULT 1
AF276872
LOCUS 1743 bp mRNA linear ROD 28-FEB-2001
DEFINITION Mus musculus sodium and chloride-dependent high-affinity choline transporter mRNA, complete cds.

ACCESSION AF276872 GI:13162669
VERSION AF276872.2
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 1743)
REFERENCE
AUTHORS Apparsundaram, S., Ferguson, S.M. and Blakely, R.D.
TITLE Molecular cloning and characterization of human and murine high-affinity choline transporters

JOURNAL unpublished
REFERENCE
AUTHORS 2 (bases 1 to 1743)
TITLE Apparsundaram, S., Ferguson, S.M. and Blakely, R.D.
JOURNAL Direct Submission
REFERENCE
AUTHORS Submitted (09-JUN-2000) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at Pierce, Nashville, TN 37232-6420, USA
REFERENCE
AUTHORS 3 (bases 1 to 1743)
TITLE Apparsundaram, S., Ferguson, S.M. and Blakely, R.D.
JOURNAL Direct Submission

JOURNAL

Submitted (28-FEB-2001) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at

Pierce, Nashville, TN 37232-6420, USA

REMARK

Sequence update by submitter

COMMENT On Feb 28, 2001 this sequence version replaced gi:11527247.

FEATURES

Source

Location/Qualifiers

1..1743

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

1..1743

/codon_start=1

/product="sodium and chloride-dependent high-affinity

choline transporter"

/protein_id="AAG36945.2"

/db_xref="gi:11527247"

/translation="MPFHEGLVATILFYLILFLVGIWAMKTKNSGPEERSEALIV

GGRODILVGGFTMTATWVGGSYINGTAEAVYGGCGIAMAAPIGYLSLILGLFP

AKPRMSKVTMTMDPFOKITGKRMGLFIPALMEEMRMAAIPESALGATTSITIDV

VNIIVTVALIAILTYLVGLSVATVDVQLFCIFIGLWISVPALSHPATDIGFT

AVHAKYQSPWLGTEISVEVYTWLDFLMLGIPMOYFORWISSATYAQVVISFL

AAFGLVVALPAICIGAIQASIDMNTAYGYDPKTKESADMLPIVQYICPVVISF

FGLGAVSAVMSADSSIIISASMFARINYOISFRONSDKEIWMVMTITVFCASA

TAMLLTKTVYGLWYLSDDIVYIIIPQLICVLFKGNITGA VAGYIEGLRTTG

EPIYLQPLIPIPGYISDKNGYINORPFKTIISMTSPFINI CVSYIAKYLPESGTLR

FKLDVFDVAVHSHSENMKTLIVRENENIKMLNPVAKROSLLTSTFTNKEALLDV

DSPSPGSGTEEDNLQ"

ORIGIN

Alignment Scores:

Pred. No.: 0
Score: 2993.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%

Length: 1743
Matches: 580
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-724-806-4 (1-580) x AF276672 (1-1743)

1 MetProphHisValGluGlyLeuValAlaIleLeuPheTyrLeuLeuIlePheLeu 20
1 ATGCTTTCCATGCTGGAAGGACTGGAGCTATTAATCTCTTCTAATCTCTTATTTTCG 60
21 ValGlyIleTrrPalAlaIleTrrPheTrrLysSerGlyAsnProGluGluArgSerGlu 40
61 GTTGGAAATGCGCTGCATGAGAAACCAAAACAGGGCAACCCAGAAAGCGACGTGAA 120
41 AlaIleIleValGlyGlyValArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
121 GCCATCATATGCTGGGGCCGCTGACATTTGTTGTTGGTGGTTTTCATGACAGCC 180
61 ThrTrrPalGlyGlyGlyTrrIleAsnGlyThrAlaGluAlaValTrrGlyProGlyCys 80
181 ACCTGGGTTGGAGGAGGCTACATCATGAGCAGAGCAGAGCAGTGTATGGCCAGGTTGT 240
81 GlyLeuAlaTrrPalAlaHisAlaProIleGlyTrrSerLeuSerLeuIleLeuGlyGlyLeu 100
241 GGTCTAGCTTGGGCTCATGACCCATGGAATTTCTGAGCTCAATTTTATAGGTGCTCG 300
101 PhePheAlaLysProMetArgSerLysGlyTrrValThrMetLeuAspProPheLysGln 120
301 TTTTTCGGAAACCTATGCTTCCAGAGATATGTGACTATGTATGACCCATTCACAAACG 360
121 IleTrrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLeuMetPhe 140
361 ATCTATGAAAGCGATGGGTGGCTCTTCATCCCTGCACCTGATGGGAGAGATGTTTC 420
141 TrrPalAlaAlaIlePheSerAlaLeuGlyAlaTrrIleSerValIleIleLysPheAlaP 160
421 TGGGTGAGCAATTTTCTGCTATTAAGGGCCACCATCAGGCTATCATTAATGAGAT 480
161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180

481 GTGAACATATGGTCATTTGCTCTGCACTCATTTGCCATTTCTTATACCTTACTGGCTGG 540
181 LeuTrrSerValAlaIleTrrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrrP 200
541 CTCATCTGTGGAGATATACGATAGATGTTTCCAGCATATCTGCATTTTATATGAGACTGG 600
201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
601 ATCATGTCCTCTTGGCCCTGCTCATCTCCAGCAGCAGCATCGATTCACACCTGTG 660
221 HisAlaLysTrrGlnSerProTrrPheGlyThrIleGlnSerValGluValTrrThrTrrP 240
661 CATGCTAAATACCAAGATCCCTGGCTGGGAACATGATGATGAGTCTACACCTGG 720
241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrrGlnAlaTrrPheGlnArg 260
721 CTGTATTAATTTCTGTATTAATGATGCTGGGTGAATCCCATGGCAAGCTTATCCAGAG 780
261 ValLeuSerSerSerSerAlaThrTrrAlaGlnValLeuSerPheLeuAlaAlaPheGly 280
781 GTCCCTCTTATCTCATCTCCAGCAACCTATGCTCAGGTAATGCTTCTGGCAGCTTTGG 840
281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrrP 300
841 TGCCTGGATGATGGCTCTACCCGCCATATGATAGAGCTATTTGAGCTTCCACAGACTGG 900
301 AsnGlnThrAlaTrrGlyTrrProAspProLysThrLysGlnGluAlaAspMetIleLeu 320
901 AACCACTGCTCCTACGGGTATCCAGATCCCAAGACTAAGAGAGAGACATGATTCCTC 960
321 ProIleValLeuGlnTrrLeuCysProValTrrIleSerPhePheGlyLeuGlyAlaVal 340
961 CGGATGCTTCCAGATACCTCTGCTTGTGACATCTCTTCTTGGGCTGGTGGCTGTT 1020
341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
1021 TCACTGCTGCTCATCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
361 ArgAsnIleTrrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTrrPal 380
1081 CGGAATATCTACACACTTCTCTTCCAGCAAAATGATCAGACAAAGAAATGTGGGGCTC 1140
381 MetArgIleTrrValLeuValPheGlyAlaSerAlaTrrPalMetAlaLeuLeuThrLys 400
1141 ATGAGATCACGTGCTGTGTGGACATCTGCAAGCCATGCTTGTGTGACAGAG 1200
401 ThrValTrrGlyLeuTrrTrrTrrLeuSerSerAspLeuValTrrIleIlePheProGln 420
1201 ACTGTGATGGGCTTGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTrrGlyAlaValAlaGlyTrrIle 440
1261 CTGCTCTGTGATCTTCTATCAAAAGAAACCAACTTATGGGGCAGTGGCTGTATATAT 1320
441 PheGlyLeuPheLeuArgIleThrGlyGlyGlnProTrrLeuTrrLeuGlnProLeuIle 460
1321 TTTGACATATCTCTGAGATTAATGAGAGAGCCATATCTATCTTGAGGCTTAAATC 1380
461 PheTrrProGlyTrrTrrSerAspLysAsnGlyIleTrrAsnGlnArgPheProPheLys 480
1381 TTCTACCTCGTATTAATCTGACAGAAATGATATATCAATCAGAGTCCCATTTAA 1440
481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTrrLeuAlaLysTrr 500
1441 ACTCTCTCAAGTATCACTCAATTTCTTACCAAACTTGTGTTCTTATCTAGCCAAGAT 1500
501 LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaLysArg 520
1501 CTATTTGAAAGTGAACCTTGTCTCAAAATTAAGATTAATGATGCTGTGTCCAAAG 1560
521 HisSerGlnGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
1561 CACAGTGAAGAGAAATGAGACAGCACTTCTAGTCAAAATGAATAATCAATTAAT 1620

QY 541 GtuleuAlaProVallyrProAroGInserLeuThrLeuSerSerTherPherThaenLys 560
 DB 1621 GAACTTGCACCTGTGAAACCTCGGACAGACCTTACCTTCAGTTCCACCAATTAAG 1680
 QY 561 GtulaAlaLeuAlaSpValAspValAspSerSerProGluGlySerGlyThrGluAspAenLys 580
 DB 1681 GAGGCGCCCTCTGATGTGATTCAGTCCGAGAGGGGTCTGGGACCTGACATTAATTACAA 1740

RESULT 2
 BC065089 4715 bp mRNA linear ROD 30-JUN-2004
 LOCUS Mus musculus solute carrier family 5 (choline transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.
 BC065089
 ACCESSION BC065089.1 GI:40787811
 VERSION MGC.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.
 1 (bases 1 to 4715)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huylk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y., Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 4715)
 Strausberg, R.
 Direct Submission
 Submitted (02-JAN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fieseler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trouc, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11528509.
 Location/Qualifiers

source
 1..4715
 /organism="Mus musculus"
 /mol_type="mRNA"
 /string="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:86023 IMAGE:6851099"
 /tissue_type="Brain, mouse, 13.5,14.5,16.5,17.5 dpc"
 /clone_1ib="NIH BMAP_FY0"
 /lab_host="DH10B"
 /note="vector: pYX-ASC"
 1..4715
 /gene="Slc5a7"
 /note="synonym: CHT1"
 /db_xref="GeneID:63993"
 /db_xref="MGI:1927126"
 27..1769
 /gene="Slc5a7"
 /codon_start=1
 /product="solute carrier family 5 (choline transporter), member 7"
 /protein_id="AAH65089.1"
 /db_xref="GI:40787812"
 /db_xref="GeneID:63993"
 /db_xref="MGI:1927126"
 /translation="MSFVHGGLVAIIIFYLLIFLVGIMAMKTKNSGNDEERSEALIV
 GGRDIGLVGFTMTAVGGVINGTAEAVYGPGLMAAPISGLIGLIF
 AKPMRSKGVVMDLDPFOIYGRMGGLIFPALMGEMFAAIFSLGANTIVYIDV
 VNISIVSALILVLTNGLXSVAYTDVVOVFCIEIGMISVFPALSHPAVMDIGFT
 AVHAKROSGLTSSVEYTVLDNLFLLMGITPMQAYFORVLSSTSTAYOVSFL
 AAFGLVMAIPACIGALIGASIDYMDQTAAYGDPKKEADMLPVLVLYCPVYSF
 IFGLAVSAVMSASDISASSMARNYVUSFRONASDKREIVWMTITLVFAGSA
 FMALITKIVYGLMVLSDLYVILIFPOLCVLFIKNTVYGAIVYIGLFLRTGG
 EYVLYOPLIFPYGYSDKNGLYNORFPKTSVMTVSFTNVCSTLYLAFPSGTLF
 PLVDVFDVAVHAHSEENMDKTLVNRNENIKMBELAPVKRQSLTSTTNKALLDV
 DSSPSGSTEINLQ"

ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 4715
 Score: 2973.00 Matches: 577
 Percent Similarity: 99.7% Conservative: 1
 Best Local Similarity: 99.5% Mismatches: 2
 Query Match: 99.3% Indels: 0
 DB: 6 Gaps: 0

US-10-724-806-4 (1-580) x BC065089 (1-4715)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
 DB 27 ATGCTCTTCCAGTAGAAGAGCTGTAGCTATATCTCTCTTACCTCTTATATTTCTG 86
 QY 21 ValGlyIleTyrPalaIaIaTrrPysThrIlyAsnSerGlyAspProGluGluAgsSerGlu 40
 DB 87 GTTGGAAATATGGCTGCATGAGAAACCAAAAGCGGCAACCAAGAGCGCAGGAA 146
 QY 41 AlaIleIleValGlyGlyIlyrGAspIleGlyLeuLeuValGlyGlyPheThrMetTrpAla 60
 DB 147 GGCATCATAGTCGGGGCGCGTGCATTTGGTTTGGTTGGTTTACCATGACAGCC 206
 QY 61 ThrTrpValGlyGlyGlyIlyrIleAsnGlyThrAlaGluAlaValIlyrGlyProGlyCys 80
 DB 207 ACCTGGGTGGAGAGAGCTACATCATATGGGACAGCAAGACAGTGTATGGCGCGGTTGT 266
 QY 81 GlyLeuAlaTrpAlaHisIaProlleGlyTyrSerLeuSerLeuIleLeuGlyIlyLeu 100
 DB 267 GGTCTAGCTTTGGCTCAGGCAACCCATTTGATTTCTTGAAGTTAATTTTGGTGTCTG 326
 QY 101 PhePheAlaIlyrPheMetArgSerIlyGlyTyrValThrMetLeuAspPropheLysGln 120
 DB 327 TTTTTCGCAACCTATGCTTCACAGGATATGATCATATGTTAGACCATTTCAACAG 386
 QY 121 IleTyrGlyIlyrArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyIlyMetPhe 140

```

Db      387 ATCTATGAAAGCGATGGGCTGCTCTTCATCCCTGCACGTAGTGGAGAGATGTTTC 446
Qy      141 TTPAALAAALAEPhSeSerAlaLeuGlyAlaThrIleSeSerValIleIleAspValasp 160
Db      447 TGGGCTGCAGCAATTTTCTGTCATTAAGGGGCCACATCAGGCTGATCTATTAGTGGAT 506
Qy      161 ValaenIleSeSerValIleValSeSerAlaLeuIleAlaIleuTyThrIleuValGly 180
Db      507 GTGAACAAATCGGTATGTCCTGCACATTCGACATTCCTTTAAACCTTACGCGGTGG 566
Qy      181 LeuTySeSerValAlaTyThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db      567 CTCTACTCTGGGCATTAATCTGATGTGTCCAGCTATTCTGCATTTTATTAAGACTGTGG 626
Qy      201 IleSeSerValProPheAlaLeuSeSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db      627 ATCAGTGTCCTTTTGCCCTGTGCACATCCTGCAGTCCAGCATCGGATTCACAGCTGTG 686
Qy      221 HisAlaIaTyThrGlnSeSerProTrpLeuGlyThrIleGlySeSerValGlnValTyThrTrp 240
Db      687 CATGCTAAATACCAAGATCCCTGGGAGAACCATTTGAAATCGTAAGTCACTTACAGCTGG 746
Qy      241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyThrPheGlnArg 260
Db      747 CTGTGTAATTTTCTGTATTGATGCTGGGTGGAAATCCATGCGCAAGCTTACCTTCCAGAG 806
Qy      261 ValLeuSeSerSeSerSeSerAlaThrTyAlaGlnValLeuSeSerPheLeuAlaIlePheGly 280
Db      807 GTCCCTCTTCATCCCTGCAGCCACATATGCTCAGTACTGCTCTCCCTGCGACCTTTTGGG 866
Qy      281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSeSerThrAspTrp 300
Db      867 TGCCGTGGATGAGCTCTACCCGCCAATGCAATAGAGAGCTATTGGAGCTTCCACAGACTGG 926
Qy      301 AsnGlnThrAlaTyGlyTyProAspProLysThrLysGlnGlyAlaIleAspMetIleLeu 320
Db      927 AACCAAGACTGCCTACGGGTATCCAGATCCCAAGATCAAGAGAGAACGACATGATTC 986
Qy      321 ProIleValLeuGlnTyLeuCyProValTyIleSeSerPhePheGlyLeuGlyAlaVal 340
Db      987 CCGATCGTTCGACATCCTGCGCCTGTGATCATCTCTTGGGCTGTGCTGCTGTT 1046
Qy      341 SerAlaAlaValMetSeSerSerAlaAspSeSerIleLeuSeSerAlaSeSerMetPheAla 360
Db      1047 TCAGCTGCTGTATGCTCCACGCTGATCTGTCATCTGTGCGAGATTCATAGTTTGGT 1106
Qy      361 ArgAsnIleTyGlnLeuSeSerPheArgGlnAsnAlaSeSerAspLysGlnIleValTTPVal 380
Db      1107 CGGAATATCTACACAGCTTCCCTTCAGACAAATGCATCGACAGAAATGTGTGGTTC 1166
Qy      381 MetArgIleThrValLeuValPheGlyAlaSerAlaIleThrAlaMetAlaLeuLeuThrLys 400
Db      1167 ATGAGGATCATCTGCTGTGTGTCGAGCATCGCAACGACATGGCTTGTGTCAGCAAG 1226
Qy      401 ThrValTyThrGlyLeuTyThrIleuSeSerAspLeuValTyIleIleIlePheProGln 420
Db      1227 ACTGTGATAGGCTCTGTGACTGAGCTTGACTTGCTTGAATCAATCACTTCCACAG 1286
Qy      421 LeuLeuCyValLeuPheIleLysGlyThrAsnThrTyGlyAlaValAlaGlyTyIle 440
Db      1287 CTGCTCTGTGATCTTTCATCAAGGAACCAACACTTAAGGGGCACTTGTGCTGTTATAT 1346
Qy      441 PheGlyLeuPheLeuArgIleThrGlyGlyIleProTyLeuTyThrLeuGlnProLeuIle 460
Db      1347 TTTGACATATTCCTAGATTAATCTGAGAGAGAGCCATATCTAATCTTACCTTAAATC 1406
Qy      461 PheTyTrpProGlyTyThrSeSerAspLysAsnGlyIleTyAsnGlnArgPheProPheLys 480
Db      1407 TTCTACCTGCTGTTATCTCTACACAAATGGTATTAACAATCAGAGGTTCCATTTAA 1466
Qy      481 ThrLeuSeSerMetValThrSeSerPheThrAsnIleCysValSeSerTyIleuAlaLysTy 500
Db      1467 ACTCTCTCATGTTACCTCATTTCTTACCAACATTTGTTCTTATCTTACCAAGATAT 1526

```

```

Qy      501 LeuPheGlnSeSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520
Db      1527 CTATTGAAAGTGAACCTTGCTCCCAAAATTAATGATGATTTGATGCTGTGTGCGCAAG 1586
Qy      521 HisSerGlnGlnAsnMetAspLysThrIleLeuValAlaGlnGlnAsnIleLysLeuAsn 540
Db      1587 CACAGTGAAGAGAAATGAGAACCAACATTCAGTCAGAAATGAAATATCAATTAAT 1646
Qy      541 GluLeuAlaProValIleProArgGlnSeSerLeuThrLeuSeSerThrPheThrAsnLys 560
Db      1647 GAAGCTGCACTGTGAAGAACTCGGAGAGCTTAACCTCAGTCACTTACCAATTAAG 1706
Qy      561 GluAlaLeuLeuAspValAlaAspSeSerProGlnGlySeSerGlyTyThrGluAspAsnLeuGln 580
Db      1707 GAGGCTCTCTGATGATGATTCACGTCGAGAGGGGCTGCGAGACTGAAGATTAATCA 1766

RESULT 3
BD012720
LOCUS      BD012720          1743 bp      DNA      linear      PAT 02-AUG-2002
DEFINITION High-affinity choline transporter.
ACCESSION  BD012720
VERSION    BD012720.1 GI:22092909
KEYWORDS   WO 0116315-A/4.
SOURCE
ORGANISM   Mus musculus (house mouse)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1743)
            Haga, T. and Okuda, T.
            High-affinity choline transporter
            Patent: WO 0116315-A 4 08-MAR-2001;
            JOURNAL
            JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA
COMMENT
OS   Mus musculus (mouse)
PN   WO 0116315-A/4
PD   08-MAR-2001
PF   18-AUG-2000 WO 2000JP005545
PR   27-AUG-1999 JP 99P 240642, 27-DEC-1999 JP 99P 366991 PI
TATSUYA HAGA, TAKASHI OKUDA
PC   C12N15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC
    A61K38/17,
    A61K45/00, A61P25/28, G01N33/53, A01K67/027

FEATURES
            source
            FT      CDS
            Location/Qualifiers
            1..1743
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      1743
Score:          2968.00      Matches:      576
Percent Similarity: 99.5%      Conservative: 1
Best Local Similarity: 99.3%      Mismatches: 3
Query Match:      99.2%      Indels: 0
DB:              2      Gaps: 0

US-10-724-806-4 (1-580) x BD012720 (1-1743)
Qy      1 MetProPheHisValGlnGlyLeuValAlaIleIleLeuPheTyThrLeuIlePheLeu 20
Db      1 ATGTCTTCCAGCTGAGAGACTGTGATCTATATCTCTTCTAATCTTATATTTCTG 60
Qy      21 ValGlyIleTTPAAlaIleTTPLyThrLysAsnSerGlyAsnProGlnGluArgSerGlu 40
Db      61 GTTGAATATGCGCGCATGAGAAACCAAAAACAGGGCAACCCGAAAGACACGTGA 120
Qy      41 AlaIleIleValGlyIleArgAspIleGlyLeuLeuValGlyIlePheThrMetThrAla 60

```


JOURNAL Submitted (14-AUG-2000) Brnesh M., University of Bonn, Pharmacology
and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY
FEATURES Location/Qualifiers
source

/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/cJ"
/db_xref="taxon:10090"
/tissue_type="Brainstem"
1..1743
/gene="CHT1"
1..1743
/gene="CHT1"
/function="sodium- and chloride-dependent reuptake of
chooline"
/experiment="experimental evidence, no additional details
recorded"
/codon_start=1
/product="high affinity choline transporter"
/protein_id="CAC03719.1"
/db_xref="GI:9843809"
/db_xref="GOA:Q9BSW5"
/db_xref="UniProtKB/TrEMBL:Q9BSW5"
/translation="MSFVEGHLVALIILFYILFLVGIWAAMTKNSGNPEHSEALIV
GARDIGLVGFTMTATVWGCGYINGTAVAVYGGCGLAMQAQPGYLSLIGLFP
AKPMRSKVTMTLDPFOQIYGRMGGLPIPALMGEMTRAAIFESALGTTISVITD
VNISYVSALAILLYLVGLYSAVTVVOLFCIFIGLWISVPALSHPAVDTIGFT
AVHAKYOSPMWLTIESVEVYTWLDFLLMLGIPMOAYFORVLSSTAYAAQVLSYL
AAGCLVVALPAICGAIASIDNMQNTAGYDPDKTEADMLIPVLYQYCPVYSF
FGLGVAASVMSADSILISASMEFARNIYOLSPRONSKDEIVWMVMTIVVFCASA
TAMALTITVVGAWLSSDLYIILIPOLCYLPIKGNVYGAVAGYIEGLFIRITGG
EPYLILOPIPIPGYSDKNGIYNORPDKTISMTSPNTICVSYLAKYLFESSTLP
FKLDVFDVAVARHSEENMDKITLVRENINKNELAPVKEROSTLUSFTFNKALDV
DSSPBGSGTEIDLQ"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 1743
Score: 2958.00 Matches: 574
Percent Similarity: 99.3% Conservative: 2
Best Local Similarity: 99.0% Mismatches: 4
Query Match: 98.8% Indels: 0
DB: 6 Gaps: 0

US-10-724-806-4 (1-580) x MM0401467 (1-1743)

QY 1 MetProPheHISValIGLUGLYLeuValAlaIleIleLeuPheTYrLeuLeuIlePheLeu 20
DB 1 ATGCTTTCCAGTGAAGGAGCTGGTAGCTATTATCTCTTCTA CTTCTATATTCTG 60
QY 21 ValGlyIleTPAlaIleATrPlySThrLYSAenSerGlyYAsnProGluGluArgSerGlu 40
DB 61 GTTGAATATGGGCTGCATGGAACCAAAAACAGGGGCAACCCAGAAAGACAGCTGAA 120
QY 41 AlaIleIleValIGLYLYArgAspIleGlyLeuLeuValIGLYLPhethrMetThraIa 60
DB 121 GCCATCATAGTGGGGGCCCTGACATGGTTGTTGGTGGTTTATTCATGACGAGCC 180
QY 61 ThrTPValIGLYLYTYrIleasnGlyThraIaGluAlaValTYrGlyProGlyCys 80
DB 181 ACCTGGTTGGAGAGGCTACATCAATGGGACAGCTAGCTGATGGGCAAGTTT 240
QY 81 GlyLeuAlaITrPAlaHISAlaProIleGlyTYrSerLeuSerLeuIleLeuGlyLeu 100
DB 241 GGTCTAGCTGGGCTCAGGCAACCATTTGATATCTCTAGCTAATTTTAAAGTGGCTG 300
QY 101 PhePheAlaLYsProMetArgSerLYsGlyTYrValIThrMetLeuAspProPheLYsGln 120
DB 301 TTTTTCGGAAACCTATGCGTTCCAAAGGATATGACATATGTTACACCATTTCAACG 360
QY 121 IleTYrGlyLYsArgMetGlyLYsLeuLeuPheIleProAlaLeuMetGlyLYsIleMetPhe 140
DB 361 ATCTATGAAAGCGATGGGTGGCTGCTCTTCACTCGACCTGATGGGAGAGATGTTTC 420

QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGCAGAAATTTTCTGTCATTTAGGGGCCAACCATCAGCGATCATTTGATGTGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTYrThrLeuValIGLY 180
DB 481 GTGAACATATGCGATATGTCTCCGACCTCATTCGCAATCTTTATACCTTATGAGGTGG 540
QY 181 LeuTYrSerValAlaTYrThrAspValaGlnLeuPheCysIlePheIleGlyLeuTrp 200
DB 541 CTCTACTGTGGGATATATGATGATGTTCCAGCATTTCTGATTTTATAGACTGTG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValIThrAspIleGlyPheThraIaVal 220
DB 601 ATCATGTCTCCCTTTTGGCCCTGTCAATCTCAGACACGACATCGAATTCACACCTGTG 660
QY 221 HISAlaLYsTYrGlnSerProTrpLeuGlyThrIleGlnSerValaGluValTYrThrTrp 240
DB 661 CATGCTAAATACCAAGATCCCTGGCTGGGAACCATTTGAATCATGTTGAAGCTTACACCTG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyLYsIleProTrpGlnAlaTYrPheGlnArg 260
DB 721 CTGTATATTTTCTGTATTGATGCTGGGTGAATCCCATGGCAAGCCTTACAGAGG 780
QY 261 ValLeuSerSerSerSerAlaThrTYrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 781 GTCCCTCTTCACTCTCAGCAACCATTCAGGATACGTCTCACTTGGCAGCTTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThraspTrp 300
DB 841 TGCCGTGATAGCTCTTACCCGCAATATGATAGAGCTATTGAGCTTCCACAGACTGG 900
QY 301 AsnGlnThraIaTYrGlyTYrProAspProLYsThrLYsGlnGluAlaAspMetIleLeu 320
DB 901 AACCAAGCTGCCTACGGGTATCCAGATCCAAAGTAAGAGAGAGACAGATGATTC 960
QY 321 ProIleValLeuGlnTYrLeuCyserProAlaTYrIleSerPhePheGlyLeuGlyAlaVal 340
DB 961 CGGATCTTCTGCACATACCTTCCCTGTGTATGACATCTCTTCTTGGGCTGGTGTCT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1021 TCAGCTGCTGTCATGTCCTTCAAGCTGACTCTGCCATCTCTGGGCAAGTTCTATGTTGCT 1080
QY 361 ArgAsnIleTYrGlnLeuSerPheArgGlnAsnAlaSerAspLYsGlnIleValITrPAla 380
DB 1081 CGGAATATCTACAGCTTCTCTTCCAGCAAAATGCATCAGACAAAGAAATTTGTGGGCTC 1140
QY 381 MetArgIleThraIleValPheGlyAlaSerAlaThraIaMetAlaLeuLeuThraIys 400
DB 1141 ATGAGGATCACGTGCTGTGTGTGGAGCATCTGCAAGCCATGCGCTTGTGTGACGAAG 1200
QY 401 ThrValTYrGlyLeuTrpTYrLeuSerSerAspLeuValTYrIleIleIlePheProGln 420
DB 1201 ACTGTGATGGGCTGTGATCTGATGCTGACCTGTGCTATCATCATCTTCCACAG 1260
QY 421 LeuLeuCysValLeuPheIleLYsGlyThrasnThrTYrGlyAlaValaIaGlyTYrIle 440
DB 1261 CTGCTCTGTGATCTTTCATCAAGAAACCAACTTATGGGGCAGTGTGCTGTTATATT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyLYsIleupProTYrLeuTYrLeuGlyProLeuIle 460
DB 1321 TTTGACATATTCCTGAGATTTATCTGAGAGAGCCATATCTATCTTGAAGCCCTTATTC 1380
QY 461 PheTYrProGlyTYrTYrSerAspLYsasnGlyIleTYrAsnGlnArgPheProPheLYs 480
DB 1381 TTCTATCCCTGTTATTTATCTGACAGAAATGTTATATCATTCAGAGGTTCCCATTTAAA 1440
QY 481 ThreLeuSerMetValThrSerPhePheThrasnIleCysValSerTYrLeuAlaLYsTYr 500
DB 1441 ACTCTCTCAATGTTACCTCATTTCTTACCAAACTTTGTGTTCTTATCTAGCCCAAGTAT 1500

QY 501 LeuphGluSerGlyThirLeuProProLySLeuAspValPheAspAlaValAlaArg 520
DB 1501 CTAATTTGAAGAGGAACTTGCTCCAAATTAATGATTTGATCTGTGTGCGAAG 1560
QY 521 HisSerGluGluAsnMetAspLysThrIleValAlaArgAsnGluAsnIleLysLeuAsn 540
DB 1561 CACAGTGAAGAGAACATGACACAAACCATTTCTAGTCAGAAATGAAATATCAAAATTTAAAT 1620
QY 541 GuleuAlaProValLysProArgGlnSerLeuThreSerSerThrPheThrAsnLys 560
DB 1621 GAACCTTGACCTGTGAACCTCGGACAGCCTTAACCTCACTTCACTTCAACCAATTAAG 1680
QY 561 GAlaAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAsnLeuGln 580
DB 1681 GAGGCCCTCTTGATTTGATTTCACTCCGAGGGGCTGTGGAGCTGAAGATTAATCTTAACA 1740
RESULT 7
BD012718 1743 bp DNA linear PAT 02-AUG-2002
LOCUS High-affinity choline transporter.
DEFINITION BD012718
ACCESSION BD012718.1 GI:22092907
VERSION WO 0116315-A/2.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 1743)
REFERENCE Haga, T. and Okuda, T.
AUTHORS High-affinity choline transporter
TITLE Patent: WO 0116315-A 2 08-MAR-2001;
JOURNAL JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA
COMMENT OS Rattus norvegicus (rat)
PN WO 0116315-A/2
PD 08-MAR-2001
PR 18-AUG-2000 WO 2000JP005545
PF 27-AUG-1999 JP 99P 240642, 27-DEC-1999 JP 99P 368991 PI
TATSUYA HAGA, TAKASHI OKUDA
PC C12M15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12M5/10, PC
A61K38/17,
PC A61K45/00, A61P25/28, G01N33/53, A01K67/027
CC
FH Key Location/Qualifiers
FT CDS (1). (1743).
Location/Qualifiers
1. 1743
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 1743
Score: 2956.00 Matches: 572
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 98.6% Mismatches: 0
Query Match: 98.8% Indels: 0
DB: 2 Gaps: 0
US-10-724-806-4 (1-580) x BD012718 (1-1743)
QY 1 MetProPheHisValGlyLeuValAlaIleLeuPheThyrLeuLeuIlePheLeu 20
DB 1 ATGCCCTTCATGTAGAGAGCTAGTAGCATTAATCTGTCTTCACTTTTAATATTCG 60
QY 21 ValGlyIleThrAlaIleTrpLysThrLysAsnSerGlyAsnProGluGluArgSerGlu 40
DB 61 GTTGGAAATATAGGCTGCATGAGAAACCAAAAACAGCGGTAAATGACAGAGAACCGAGAA 120
QY 41 AlaIleIleValAlaGlyLysArgAspIleGlyLeuLeuValGlyLysPheThrMetThrAla 60
DB 121 GCCATCATGATGTGGGGCCGAGACATTTGTTGTGGTGGTGTATTACATGACAGCC 180

QY 61 ThrTrpValGlyGlyGlyThrIleAsnGlyThrAlaGluAlaValTrpGlyProGlyCys 80
DB 181 ACTGGGTTGGAGAGAGGTACATCAACAGGAGACGTCGAAAGCATTTATGGGCCAGGTTGT 240
QY 81 GlyLeuAlaIleThrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyLeu 100
DB 241 GGTCTAGCTTGGGCTCAGGACCCCATTTGATTTCTGAGCTCTGATTTAGGTGGCCCTG 300
QY 101 PhePheAlaLysProMetLysArgSerLysGlyTyrValThreLeuAspProPheLysGln 120
DB 301 TTTTGGCAAAACCTATAGGCTTCAGAGGATATGACATATGTACACCGCTTCAACAG 360
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyLysMetPhe 140
DB 361 ATCTATGAAAGCGCATGGGTGGGCTGCTTCACTCCGCACTGATGGAGAGAGTTC 420
QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGACAGCAATTTTCTCTGATAGGGGCTACCATCAGCGTAATCATTTGATGTGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuThyrThreLeuValGly 180
DB 481 GTGAACATATCGGTCAATGTCTCCGACATTCATTCATTTATACCTCGTGGAGAGG 540
QY 181 LeuTyrSerValAlaIleThyrThraPValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
DB 541 CTCTACTCTGTGGCATATATCTGATGTGTACACTTTCATCTTTTAATAGATTTGTGG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB 601 ATCAGTGTCCCATTTGCCCTGTCACATCTGCAGTCACGACATTCATGATTCACCTGCTG 660
QY 221 HisAlaLysTyrGlnSerProTrpLeuGlyThrIleGluSerValGluValTyrTrpTrp 240
DB 661 CATGTAAATATACAGAGTCTCTGGGAGAACCATTAAGTCAAGCTCAACCTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrPheGlnArg 260
DB 721 CTTGATTAATTTCTGTGTGTGTGATGCTGGGTGAATACATGGAGACCTTCCAGAGG 780
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 781 GTCCTCTCTTCATCGACAGACCTATGCTCAGGTCTGCTCTCGGACGCTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyLysSerThrAspTrp 300
DB 841 TGCCTGTGATGCTCTTACAGCCATTTGCATTTGGGGCCATTTGGACCTTCCACAGACTGG 900
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetIleLeu 320
DB 901 AACCAACCTGCATATAGGCTTCCAGATCCCAAGACCAAGAGAGACATCATTTCTC 960
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
DB 961 CCGATTTGTTCTACAGTACCTGCTGCTGATCATTTCTTGTGGGCTGTGCTGCTT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerMetPheAla 360
DB 1021 TCTGCTGCTGATAGTCCCTGCTGCTGATCATTCATTCACAGAGTTCCATGTTTGTCT 1080
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrpVal 380
DB 1081 CGGAATATCTACAGCTTCTTCTTCAACAATAAGCATCAGACAGAAATGTGTGGTCTC 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
DB 1141 ATGAGGATCATCTGTGTTGTGTGGAGCATCGCAACGACCATGCTGTGCTCAGAG 1200
QY 401 ThrValTyrGlyLeuThrTrpTyrLeuSerSerAspLeuValTyrIleIlePheProGln 420
DB 1201 ACTGTATATGGGCTCTGTGATCTGAGCTCTGACCTTGTCTACATCATCTTTCCACAG 1260

Db 1021 TCTGCTGCTGTCATGTCCTCGCTGACTCATCTCATCAGCAAGTTCATGTTGCT 1080
Qy 361 ArgAsnIleTyrGlnIleuSerPheArgIAsnAlaSerAspIysGluIleValTPrpVal 380
Db 1081 CGGAATATCTACAGAGCTTCTTCAGCAAAATGATCGAGCAAGGAATATGTGGGTC 1140
Qy 381 MetArgIleThrValIleuValPheGlyAlaSerAlaThrAlaMetAlaIleuLeuThrIys 400
Db 1141 ATGAGATCATCTGTGTTGTGTGGAGCATTCGCAACGCCATGGCTTGCTCAGGAAG 1200
Qy 401 ThrValIleGlyLeuTPrpIleuSerSerAspIleuValTyrIleIleIlePheProGln 420
Db 1201 ACTGTATATGGGCTCGAGTACCTGAGCTGCTGTCTACATCATCTCCACAG 1260
Qy 421 LeuLeuCyValIleuPheIleIysGlyTPrpAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db 1261 CTGCTCTGTGATCTTCATCAAGGAACCAACATATGGGCGAGTGTGGTATAT 1320
Qy 441 PheGlyLeuPheLeuArgIleThrGlyGlyLeuProTyrLeuTyrIleuGlnProLeuIle 460
Db 1321 TTTGGACTTTTCTCGAATATACCGAGAGGAGCCATATCTATACCTTGACGCCCTTAATC 1380
Qy 461 PheTyrProGlyTyrTyrSerAspIysAsnGlyIleTyrAsnGlnArgPheProPheIys 480
Db 1381 TTCTACCTGCTATTACCTCGCAAGAAATGATATACATACAGAGGTTCCCATTTAAA 1440
Qy 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCyValSerTyrLeuAlaIysTyr 500
Db 1441 ACTCTCTCAATGGTATACCTCATCTTATACCAATTTGTTTCATATCAGCAAGTAT 1500
Qy 501 LeuPheGluSerGlyThrLeuProTyrLeuAspValPheAspAlaValAlaIAsn 520
Db 1501 CATTTGAAAGTGGACCTTGCTCCAAATATGATATATGATCTGTGTCTCAAG 1560
Qy 521 HisSerGluGluAsnMetAspIysThrIleLeuValArgAsnGluAsnIleIysLeuAsn 540
Db 1561 CACAGTGAAGAGAACATGCGACAGACCAATCTAGTCAGAAATGAAAAATCAATTTAAAT 1620
Qy 541 GlnLeuAlaProValIysProArgIleuSerLeuThrLeuSerSerThrPheThrAsnIys 560
Db 1621 GAATCTGACCTGTAAAGCTCGACAGAGCTTACCTCATCTTCACTTCACTTCACTTAA 1680
Qy 561 GlnAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAsnLeuGln 580
Db 1681 GAGGCTCTCTTGTATTTGATTTCACTCCAGAGGATCTGGGACTGAGATTACTTACA 1740

RESULT 9
AB030947 4904 bp mRNA linear ROD 03-FEB-2000
LOCUS Rattus norvegicus mRNA for high-affinity choline transporter CHT1,
DEFINITION complete cds.
ACCESSION AB030947
VERSION AB030947.1 GI:6863033
KEYWORDS choline transporter; high-affinity choline transporter CHT1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Scleroglossa; Muridae; Muridae; Murinae; Rattus.
1 (sites)
Okuda, T., Haga, T., Kanai, Y., Endou, H., Ishihara, T. and Katsura, I.
Identification and characterization of the high-affinity choline
transporter
Nat. Neurosci. 3 (2), 120-125 (2000)
JOURNAL PubMed 10649566
AUTHORS Okuda, T.
TITLE Direct Submission
SUBMITTED (09-AUG-1999) Takashi Okuda, University of Tokyo, Faculty
of Medicine, Department of Neurochemistry, Hongo 7-3-1, Bunkyo-ku
113-0033, Japan (E-mail: okuda@u-tokyo.ac.jp, Tel: +81-3-5841-3560,
Fax: +81-3-3814-8154)
COMMENT Sequence updated (11-Jan-2000).

FEATURES
source Location/Qualifiers
1..4904
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/clone="CHT1"
/tissue_type="spinal cord"
/clone_lib="rat spinal cord cDNA library"
/dev_stage="adult"
224..1966
/codon_start=1
/product="high-affinity choline transporter CHT1"
/protein_id="BAA90484.1"
/db_xref="GI:6863034"
/translation="MPFHEGVALIIEFLIIEFLICIAAMAKTKNSGAESEATIV
GGRDGLVGFTRATWVGGINATBAVVGPGGLMAQAPISLILGILF
AKPMSKGVYVMDLPQOIVGRMGLEFIPALMEFWAAAI FSLGATISVILIDV
VNIISVIALIILYLVGLYSVAVTDVOLFIEIGLISVPELSPAVDILGFT
AVHAKVSPMLGTIESVEYVWLDNLFILMLGIPQOAFORVLSSTAYOVLSFL
AARGCLVMAIPACIGATGASTDMOTAVGPDPPKREADMILPILQYLCVYISF
PGLGVAASVMSADSSITLSASMTARNTYQSFQNSADKELVWMTITVPEGSA
TMALITKTIVGLWYLSDDVITIIIFPOLCVLFGKNTTGVAGVITGLERKRTGG
EPYLQPLIFPYGYPDPKNGIYNORFPKRLTSMYTSFTNLCVSYLARYLESGLP
PKLIDFVAVSHSEEMDKTILVRENINKLIELAPVKPROSLTUSSTFTNKELIDV
DSSPGSGTEEDMLQ"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4904
Score: 2956.00 Matches: 572
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 98.6% Mismatches: 4
Query Match: 98.8% Indels: 0
Gaps: 0

US-10-724-806-4 (1-580) x AB030947 (1-4904)

Qy 1 MetProPheHisValGluIleuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 224 ATGCTCTTCCATGTAAGAGCTGAGTATATCTGTCTTCACTTCTTATATTTCTG 283
Qy 21 ValGlyIleTPrpAlaIleTPrpIysThrIysAsnSerGlyAsnProGluIAsnArgSerGlu 40
Db 284 GTTGGATATGGGCTGCAATGAAACCAAAACAGCGGTATATGACAAAGACGAGAA 343
Qy 41 AlaIleIleValIleGlyIleValAspIleGlyLeuLeuValIleGlyPheThrMetThrAla 60
Db 344 GCCATCATATGTTGGGGCCGAGACATTTGTTGTTGTTGTTTACCATGACAGACC 403
Qy 61 ThrTPrpValIleGlyIleTyrIleAsnGlyThrAlaGluAlaValIleTyrGlyProGlyIys 80
Db 404 ACCGTGCTGTGGAGAGGTATCAATCAACGAGACGCTGAAAGCACTTATATGGCCAGGTGT 463
Qy 81 GlyLeuAlaTPrpAlaHisAlaProIleGlyTyrSerLeuSerLeuIleuGlyGlyLeu 100
Db 464 GGTCTAGCTTGGGCTCAGGACCCATTTGATTTCTAGTCTGATTTTAAAGTGGCTG 523
Qy 101 PhePheAlaIlePheMetArgSerIysGlyTyrValThrMetLeuAspProPheIysGln 120
Db 524 TTTTTCGAAAACCTATGCTTCCAAAGGATATGTACTATGTTAGACCCGTTTCAACAG 583
Qy 121 IleTyrGlyIleArgMetIleGlyLeuLeuPheIleProAlaIleuMetGlyIleMetPhe 140
Db 584 ATCTATGAAAGGACATGGGCTGGCTGCTGTTCTATCTGCACTATGAGAGAGATCTTC 643
Qy 141 TPrpAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 644 TGGGCTGACGAAATTTCTCTCATTAAGGGCTACCATCAGCGTATCTTGTATGGAGAT 703
Qy 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValIleGly 180
Db 704 GTGAACATATCGGATCTGTCTCCGACATCATTTGCAATTTTATACCTCTGTGGAGGG 763

181 LeuTyrSerValAlaTyrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db CTCTACTCTGGGCATATATCTGATGTGTGACAGCTATCTGCAATTTTATAGATTGG 823
201 ILeSerValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThrAlaVal 220
Db ATCACTGTCACATTTGGCCCTGTCACATCTGCAGTCACGACCATTTGGATTCTGCTGG 883
221 HisAlaIleTyrGlnSerProTyrPheGlyThrIleGluSerValGluValTyrThrTrp 240
Db CATGCTAATATACCAAGTCCCTGGGAGCAATGAAATGAGTGAAGTCTACACCTGG 943
241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArg 260
Db CTGTGTAATTTCTGTGTGTGATGCTGGGTGGAAATACCATGCAAGCCCTACCTCCAGAG 1003
261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db GTCCCTCTCTCATGCTCAGCGACCTATGCTCAGGTGCTCTCTCTGGCAGCTTTGGG 1063
281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db TGCCGTGGATGGCTCTACCAACCATTTGGGAGCATTTGGAGCTCCACAGACTGG 1123
301 AsnGlnThrAlaTyrGlyTyrProAspProIleTyrGlyGlnAlaAspMetIleLeu 320
Db AACCAACATGCATATGGGTTCACATCCCAAGACCAAGAGAGCAAGCATGATTC 1183
321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db CGGATTTGTTCTACAGTACCTGCTGCTGTGATCATTTCTTTGGGCTGGTGTGT 1243
341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db TCTGCTGCTGATGTCCTCGGCTGACCTCATCTCATCTATACAGAGTTCCATGTTGCT 1303
361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspGlyGluIleValTrpVal 380
Db CGGAATATCTACCACTTCTCTTCAGACAAATGTCATCAGACAGAAATGCTGGGGTCTC 1363
381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400
Db ATGAGGATCATGCTGTTGTTGTTGAGCATCTGCAACGACCTGCTGCTCAAGAG 1423
401 ThrValTyrGlyLeuTyrPyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
Db ACTGATGATGGGCTCTGGATCTGAGCTGCACTGTGCTCATCATCATCTTCCACAG 1483
421 LeuLeuCysValLeuPheIleIleGlyTyrThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db CTGCTCTGTGATCTTCTCATCAAGAACCAACATTAAGGGGCTGCTGTATATAT 1543
441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIle 460
Db TTTGACCTTTTCTCAGAAATTCAGGAGAGGCCATATCTATCTGACGCCCTTATC 1603
461 PheTyrProGlyTyrTyrSerAspIleAsnGlyIleTyrAsnGlnArgPheProPheIle 480
Db TTCTACCTGGTATTATACCTGACAAATGATATATACATCAGAGGTTCCCATTTAA 1663
481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaIleTyr 500
Db ACTCTCTCATAGTTACCTCATCTTATCAACATTTGTTGTTCTTATCTACCAAGAT 1723
501 LeuPheGluSerGlyThrLeuProProIleAspValPheAspAlaValAlaIleArg 520
Db CTATTGAAAGTGGAACTTGGCTCCAAATATAGATATTTGAAGCTGTTGTCTCAGG 1783
521 HisSerGluGluAsnMetAspIleThrIleLeuValArgAsnGluAsnIleIleLeuAsn 540
Db CACAGTGAAGAGAAACATGAGCAAGACATTTCTAGTCAGAAATGAAAAACATCAATTAAT 1843

541 GluLeuAlaProValIleProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIle 560
Db GAACCTTGACACCTGTAAGCCCTCGACAGGCCCTAACCCCTCACTTCAACCAATTA 1903
561 GluAlaLeuLeuAspValAlaSerSerProGluIleGlySerGlyTyrGlnAspAsnLeuGln 580
Db GAGGCTCTCTTGATGTTGATTCAGTCCAGAGGAGATCTGGGACTGAAAGATTAATTACA 1963

RESULT 10
BD012719
LOCUS BD012719 1743 bp DNA linear PAT 02-AUG-2002
DEFINITION High-affinity choline transporter.
ACCESSION BD012719.1 GI:22092908
VERSION WO 0116315-A/3.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1743)
AUTHORS Haga, T. and Okuda, T.
TITLE High-affinity choline transporter
JOURNAL Patent: WO 0116315-A 3 08-MAR-2001;
JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA
COMMENT OS Homo sapiens (human)
PN WO 0116315-A/3
PD 08-MAR-2001
PF 18-AUG-2000 WO 2000JP005545
PR 27-AUG-1999 JP 99P 240642, 27-DEC-1999 JP 99P 368991 PI
TATSUYA HAGA, TAKASHI OKUDA
PC C12N15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC
A61K38/17,
PC A61K45/00, A61P25/28, G01N33/53, A01K67/027
CC
FH key Location/Qualifiers
FT CDS Location/Qualifiers
1..1743
1..1743
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 1743
Score: 2795.00 Matches: 537
Percent Similarity: 96.64 Conservative: 23
Best Local Similarity: 92.64 Mismatches: 20
Query Match: 93.48 Indels: 0
DB: 2 Gaps: 0

US-10-724-806-4 (1-580) x BD012719 (1-1743)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 1 ATGGCTTTCATGAGGAGGACTGATGATCATGCTGCTCATCTTAATTTTGTG 60
QY 21 ValGlyIleTyrPalaIleTyrPheThrIleAsnSerGlyAsnProGluGluArgSerGlu 40
Db 61 GTTGAATATATGGGCTGCTGAGAACCAAAACAGTGCACGCGAGAGAGAGCGGAA 120
QY 41 AlaIleIleValIleGlyIleArgAspIleGlyLeuLeuValIleGlyPheThrMetThrAla 60
Db 121 GCCATCATAGTTGGTGGCCGAGATATTGTTATGTTGTTGATTTACCATGACACT 180
QY 61 ThrTyrPalaIleGlyIleTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyIle 80
Db 181 ACTGGGTGAGAGAGGGTATATCAATGAGCAGCTGAAGAGTTATGATCAAGTTAT 240
QY 81 GlyLeuAlaTyrPalaIleAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
Db 241 GGCTAGCTTGGGCTCAGGCAACCAATTGGATTAATCTTAACTTGAATTTAGTGGCTG 300


```

Db 181 ACCTGGGCGGAGGGGATATATCAATGACACAGCTGAAAGCAGTTATATCAACAGTTAT 240
Qy GlyLeuAlaTPAlaHleAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
Db 241 GGCCAGCTTGGGCTCAGGACCAATGGATATCTCTTAGCTGATTTTAGGTGGCCCTG 300
Qy PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln 120
Db 301 TTCCTTGCAGAAACCATATCGTTCAAGGGGTATGACACATGTATAGACCCGTTTCAGCAA 360
Qy IleTyrGlyLysArgMetGlyGlyLeuPheIleProAlaLeuMetGlyGlyMetPhe 140
Db 361 ATCTATGAGAAAACGATGGCGGACCTCGTTTATTTCTGCACATGATGGAGAAATGTTTC 420
Qy TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 421 TGGGCTGCAGCAATTTCTCTGCTTGGGAGCCACCATCAGCTGATCATCGATGGAT 480
Qy ValAlaAlaIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
Db 481 ATGCACATTTCTGTCATCATCTCTGCACCTATTCACACTGCTGACACTGGTGGAGGG 540
Qy LeuTyrSerValAlaIleTyrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTyr 200
Db 541 CTCTATCTGTCGCTTACCTGATGATGCTGCTGCTTTCGATTTTTCAGCTTTTTCAGGCTG 600
Qy IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db 601 ATCAGCGTCCCTTTGCAATGTGCATCTGTCAGCTGTCAGATCGGAGATCTGCTG 660
Qy HisAlaLysTyrGlnSerProTyrLeuGlyThrIleGlyLeuSerValGlyValTyrThrTyr 240
Db 661 CATGCAATATCCAAAGCCGTCGGGACCTGTGACTGCTCATCTGCAAGTCTACCTTGG 720
Qy LeuAspAsnPheLeuLeuLeuMetLeuGlyIleIleProTyrGlnAlaTyrPheGlnArg 260
Db 721 CTTCATAGTTTCTCTGTGTTGATGCTGGGTGGAAATCCCATGGCAAGCACTTTCAGAG 780
Qy ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db 781 GTTCTCTCTTCTTCTTCAGCCACCATATGCTCAAGTCTGCTCTTCTGCGACCTTTCGG 840
Qy CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaIleSerThrAspTyr 300
Db 841 TGCCGCGGATGGCCATCCAGCCATACCATTTGGGGCCATTTGGAGCATCAACAACATGG 900
Qy AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGlnGluAlaAspMetIleLeu 320
Db 901 AACCGACTGCTATGAGGCTTCCAAATCCCAAGACTACAGAAAGGCGACATGATTTTA 960
Qy ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCAATGTTCTGCACTATCTGCTGCTGATATTTCTTCTTGTGCTTGTGCTGACAGTT 1020
Qy SerAlaAlaValMetSerSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1021 TCTGCTGCTGATATGATCATACAGCAATTCCTTCATCTGTCGCAAGTCCATGTTTGA 1080
Qy ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTyrVal 380
Db 1081 CGGAACATCTACACCTTCTCTTCAGACAAATGCTTCGCAAGAAAGAAATGCTTGGGTT 1140
Qy MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db 1141 ATGCAATCAACAGTGTGTTGTGTTGAGACATCTGCAGAACGCGATGCTGTGACGAAA 1200
Qy ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
Db 1201 ACTGATGATGGCTCTGGACTCATGAGTCTGACCTGTTTACATGATATCTTCCCCAG 1260
Qy LeuLeuCysValLeuPheIleLysGlyThrAsnTyrTyrGlyAlaValAlaGlyTyrIle 440

```

```

Db 1261 CTGCTTGTGATCTTGTGTTAAGGAAACAACATATGAGGCCGTCGAGTTANGTT 1320
Qy PheGlyLeuPheLeuArgIleThrGlyGlyGlnProTyrLeuTyrLeuGlnProLeuIle 460
Db 1321 TCTGGCCCTTCTCTGAGAAATACGAGGAGGAGCCATATCTGTATCTTCAACCCCTGATC 1380
Qy PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
Db 1381 TTCATACCTGCTATTAACCCGATGATATATGATATATATATCAGAAATTTCCATTATAA 1440
Qy ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
Db 1441 AACCTGCGATGATGATCATCTTCTTAACCAACATTTCATCTCTATCTAGCAAGTAT 1500
Qy LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaIleArg 520
Db 1501 CTATTGGAAGTGGAGACCTTCCACCTAAATTAGATGATTTGATGCTGTTGCCAGAA 1560
Qy HisSerGlnGluAsnMetAspLysThrIleLeuValArgAsnGlnAsnIleLysLeuAsn 540
Db 1561 CACAGTGAAGAAAACATGATATAGCAATCTTGTCAAAAATGAAATATTAATTAAGAT 1620
Qy GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db 1621 GAACCTGCACTTGTGAGACCCAGACAGACATGACCTCAGCTCACTTCAACATATAA 1680
Qy GluAlaLeuLeuAspValAspSerSerProGlnGlySerGlyThrGlnLysAsnLeuGln 580
Db 1681 GAGGCTTCTTGTGATGTTGATTCAGTCACAGAGGCTGAGGACGTAAGATATTTACAG 1740

RESULT 12
E49871 1743 bp DNA linear PAT 27-AUG-2002
LOCUS High-affinity choline transporter.
DEFINITION E49871
ACCESSION E49871
VERSION 1
KEYWORDS JP 2001136976-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1743)
AUTHORS Haga,T. and Okuda,T.
TITLE High-affinity choline transporter
JOURNAL Patent: JP 2001136976-A 3 22-MAY-2001;
SCIENCE & TECH AGENCY
COMMENT OS Homo sapiens (human)
PN JP 2001136976-A/3
PD 22-MAY-2001
PF 27-DEC-1999 JP 1999368991
PI TATSUYA HAGA, TAKASHI OKUDA
PC C12N5/09, A01K67/027, A61K38/00, C07K14/47, C07K16/18, C07K19/00,
PC C12N5/10.
PC C12P21/02, C12P21/08, C12Q1/00, C12N15/00, A61K37/02, C12N5/00 CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
(1)..(1743).
source 1..1743
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 1743
Score: 2795.00 Matches: 537
Percent Similarity: 56.68 Conservative: 23
Best Local Similarity: 52.68 Mismatches: 20
Query Match: 93.48 Indels: 0
DB: 2 Gaps: 0

US-10-724-806-4 (1-580) x E49871 (1-1743)

```



```

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
DB 1 ATGGCTTTCCATGTGAAAGACTGATAGCTATCATCTGTTCTTACCTTTTAATTTTGGCTG 60
QY 21 ValGlyIleTyrAlaIleTyrPheTyrLeuSerGlyAsnProGluIuArgSerGlu 40
DB 61 GTTGGAAATATGGAGCTGCTGAGAAACCAAAACAGTGGCAGGACAGAGGCGACAGAA 120
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuValGlyGlyPheThrMetThrAla 60
DB 121 GCCATCATATGTTGGCCGAGATATGTTGATTTGTTGTTGATTTTACCATGACAGCT 180
QY 61 ThrTyrValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
DB 181 ACCTGGGTGGAGAGAGGATATCATATGACACAGCTGAAAGCAGTTTATGATCCAGGTTAT 240
QY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
DB 241 GGCTTAGCTTGGGCTCAGGACCAATTTGATATTTCTTAGTCTGATTTTAAAGTGGCTG 300
QY 101 PhePheAlaLysProMetLysSerLysGlyTyrValThrMetLeuAspProPheLysGln 120
DB 301 TTCTTTGCAAAACCTATGCGTTCMAAGGGGTATGTACCATGTTAGACCCGTTTCAGCAA 360
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
DB 361 ATCTATGAAAGCGATGGGCGGACTCTGTTTATTTCTGCACTGATGAGGAAATGTTTC 420
QY 141 TrrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGGTGCAGCAATTTCTCTGCTTTGGGAGCCACATCAGCGTATCATGATGATGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
DB 481 ATGCACATTTCTGTCATCTCTGTCATCTGTCATCTGTCATCTGTCATCTGTCATCTG 540
QY 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200
DB 541 CTCTATCTGAGGCTTACACTGATGTCGTTGAGCTTTTGGCAATTTTGTAGGGCTGTG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB 601 ATCAGGCTCCCTTTCATGTCATCTGTCATCTGTCATCTGTCATCTGTCATCTGTC 660
QY 221 HisAlaLysTyrGlnSerProTrrPleuGlyThrIleGluSerValGluValTyrThrTrr 240
DB 661 CATGCCAATATACCAAAAGCCGTGGCGGAACTGTGACTGTGAACTGTGAACTGTG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrrGlnAlaTyrPheGlnArg 260
DB 721 CTGTGATGTTTCTGTTGTGATGCTGGGGTGGAAATCCATGGCAACATCATCTTCAGAG 780
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 781 GTTCTCTCTCTCTTCCACGACCACTATATGTCATCAAGTCTCTCTCTTCTGACGTTT 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrr 300
DB 841 TCCCTGGTGTATGGCCATCCAGCATATCATTTGGGGCCATTTGGACATCATCAGACTG 900
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGlnGlnAlaAspMetIleLeu 320
DB 901 AACCAAGCTGCATATATGGCTTCCAGATCCACAGACTACAGAGAGGCGACATGATTTTA 960
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
DB 961 CCAATTTGTTCTGAGATCTCTGCTGCTGATATATTTCTTTTGGTCTGGTGCAGTT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerLysSerMetPheAla 360
DB 1021 TCTGCTGCTGTATATGATCATCAGCAATTTCTTCATCTGTGCAGCAAGTTCCATGTTTGA 1080

```

```

QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTrrVal 380
DB 1081 CGGACATCTTACACACTTTCTTCCAGACAAATGCTTGGACAAAGAAATCGTTGGCTT 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
DB 1141 ATGCAGATCATCAGATGTTGTTGTTGAGCATCTGCAACAGCCATGAGCTTGTCTGACGAA 1200
QY 401 ThrValTyrGlyLeuTrrPyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
DB 1201 ACTGTGATGGGCTCTGTTGACTCAGTTTCTGACTTTGTTTACATGCTTATCTTCCCCG 1260
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTrrIle 440
DB 1261 CTGCTTTGTGATCTCTTTGTTTAAAGGAACCAACACTTATGGGCGCGTGCAGGTTATG 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIle 460
DB 1321 TCTGGCTCTTCTCTGAGATATACGTGAGGGGAGCCATATCTGATCTTCCAGCCCTTGATC 1380
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
DB 1381 TTCTACCTGCTATTTACCTGATGATATGATATATATATATATATATATATATATAT 1440
QY 481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
DB 1441 ACACTTGCCATGTTATCATCATTTTACCAACCAATTTGGATCTCCATCTAGCCAAAT 1500
QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520
DB 1501 CTATTTGAAAGTGAACCTTGCACCTTAATTAATGATGATTTGATGCTCTTGTGGACA 1560
QY 521 HisSerGlnGlnAsnMetAspLysThrIleLeuValArgAsnGlnLysIleLysLeuAsn 540
DB 1561 CACAGTGAAGAAACATGATTAAGCAATTTCTTGTCAAAATGAATAATTAATTAATGAT 1620
QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
DB 1621 GAACTTGCACCTGTGTAAGCAGCAGACAGACATGACCTTACGCTCAACATTAATA 1680
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGlnAspAsnLeuGln 580
DB 1681 GAGGCTTCTCTGATGTTGATTCAGTCCAGAAAGGCTCGGAGACTGAAATTAATTACAG 1740

RESULT 13
AR268949 1743 bp DNA linear PAT 10-APR-2003
LOCUS AR268949
DEFINITION Sequence 1 from patent US 6500643.
ACCESSION AR268949
VERSION AR268949.1 GI:29699686
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1743)
AUTHORS Wu D.-H., Gu Y., Millard W.J. and He Y.-J.
TITLE Human high affinity choline transporter
JOURNAL Patent: US 6500643-A 1 31-DEC-2002;
University of Florida; Gainesville, FL
FEATURES
source
1..1743
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 0 Length: 1743
Pred. No.: 2795.00 Matches: 537
Score: 96.6% Conservative: 23
Best Similarity: 92.6% Indels: 20
Best Local Similarity: 93.4% Mismatches: 0
Query Match: 2 Gaps: 0
DB:

```

```

US-10-724-806-4 (1-580) x AR268949 (1-1743)
OY 1 MetProPheHisValGlyGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 1 ATGGCTTCCATGTCGGAAGGAGCTGATAGCTATCATCGTGTCTACCTCTAATTGTCG 60
OY 21 ValGlyIleTTPAlaAlaIleTPlYsThrLySaSerGlyAsnProGluGluArgSerGlu 40
Db 61 GTTGAAATATGGGCTGCTCGAGAACCAAAAACATGGGACGCGAAGAGCGCGCA 120
OY 41 AlaIleIleValGlyGlyYArgAspIleGlyLeuValGlyGlyPhePheMetThrAla 60
Db 121 GCCATCATAGTGGTGGCCGAGATATGGTTTATGGTTGGTGAATTTACATGACAGCT 180
OY 61 ThrTTPValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db 181 ACCTGGTCGGAGGAGGGATATCATGCAATGCAAGCTGAAGCAAGTTATGTCAGATTAT 240
OY 81 GlyLeuAlaIleTTPAlaHisAlaProIleGlyTyrSerLeuSerLeuIleGluGlyLeu 100
Db 241 GGCTTAGCTTGGGCTCAGGACCAATGGATATCTCTAGCTGATATTTAGTGGCCCTG 300
OY 101 PhePheAlaLysProMetArgSerLysGlyTyrValIleMetLeuAspProPheLysGln 120
Db 301 TTTCTTGCAAAACCTATGCGTTCAAAAGGGGTAATGACCATGTTAGACCCGTTTCAGCA 360
OY 121 IleTyrGlyLysArgMetGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140
Db 361 ATCTATGGAACACGATGGGGGCGGACCTCGTTTATTCGCACTGATGGAGAAATGTTTC 420
OY 141 TTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 421 TGGGCTGACGAAATTTCTCTGCTTGGAGCCACATCAGCGGTATCATCATGTGGAT 480
OY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
Db 481 ATGCACATTTCTGTATCATCTCTGCACTCATTTGCCACTCTGTACACTGGGGAGGG 540
OY 181 LeuTyrSerValAlaIleTyrThrAspValAlaGluLeuPheCysIlePheIleGlyLeuTyr 200
Db 541 CTCTATTCGTGGCTACACTGATGTCGTTCAGCTTTTGCATTTTGTAGGGCTGTGG 600
OY 201 IleSerValProPheAlaLeuSerHisProAlaValIleAspIleGlyPheThrAlaVal 220
Db 601 ATGACGCTCCCTTTCGATGTCACATCTCGACATCGCAGACATCGGGTCACTCTGTG 660
OY 221 HisAlaLysTyrGlnSerProTyrLeuGlyThrIleGluSerValGlyValTyrThrTyr 240
Db 661 CATGCCAAATACCAAAAGCCGTGGCTGGGAACGTGTGACTCATCTGAAGTCTACTCTTGG 720
OY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaIleTyrPheGlnArg 260
Db 721 CTGTATAGTTTCTGTGTGATGCTGGGTGGAATCCCATGCAAGCATATTTCAGAGG 780
OY 261 ValLeuSerSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db 781 GTTCTCTCTTCTTCTTCAGCCACCATATGCTCAAGAGCTGTCTCTCTCGCACCTTTCGGG 840
OY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300
Db 841 TGCCGTGGATGGCCATCCCAAGCCATCTCATTTGGGGCATTTGGAGCCCTCCACAGACTGG 900
OY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetIleLeu 320
Db 901 AACCAAGATGATGGGCTTCCAGATCCCAAGACCTACAGAGAGAGAGACATGATTTTA 960
OY 321 ProIleValLeuGlnTyrLeuGlyProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCAATGTTCTGCAATATCTCTGCTGTGATATTTCTTCTTGGTCTGTGTCAGATT 1020
OY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1021 TCTGTGCTGTTATGTCATCAGACAGATTCTTCACTTCTCAGCAAGTTCCATGTTTGA 1080

```

```

OY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValIleTTPVal 380
Db 1081 CGGAACATCTACCAAGCTTCTCTTCAGACAAATATCTTCGACAAAGAAATCGTTGGGTT 1140
OY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db 1141 ATGGCAATCACAGTGTGTTGTGGAGCATCTCAACAGCCATGCGCTTCTGACGAA 1200
OY 401 ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
Db 1201 ACTGATATGGGCTCTGCTGATCTCACTTCGATCCTTGTTTACATCGTATCTTCCCCAG 1260
OY 421 LeuLeuCysValIlePheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db 1261 CTGCTTGTGTACTTGTGTATAGGAAACCAACACTATGGGCGCGTGGCAGATTATGTT 1320
OY 441 PheGlyLeuPheLeuArgIleThrGlyGlyLysProTyrLeuTyrLeuGlnProLeuIle 460
Db 1321 TCTGGCTCTTCTCGAGAAATACAGAGGAGGACCATATCTGTATCTTCACGCCCTTGATC 1380
OY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
Db 1381 TTCTACCTGCGCTATTAACCCGATGATATGTAATATATATATATATATATATATATAT 1440
OY 481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
Db 1441 ACACCTGCGCATGTTACATCATCTTAAACCAACTTTCATCTCTCTTAACTGACCAATAT 1500
OY 501 LeuPheGlySerGlyTyrLeuProProLysLeuAspValPheAspAlaValAlaLysArg 520
Db 1501 CTATTTGAAATGGAACCTTCCACCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
OY 521 HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
Db 1561 CACAGTGAAGAAACATGATATAGACATCTTGTCAAAATATGAATAATTAATTAATTAAT 1620
OY 541 GluLeuAlaProValLysProArgLysSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db 1621 GAACCTTGACCTTGCAAGCCACGACAGACATGACCCCTCACTTCAACCAATTAATA 1680
OY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyTyrGluAspAsnLeuGln 580
Db 1681 GAGGCTTCTTGTGATGTTATTCAGTCCAGAGAGGCTCGGACTGAAGATTAATTACAA 1740

RESULT 14
AF276871 1743 bp mRNA linear PRI 27-NOV-2000
LOCUS AF276871
DEFINITION Homo sapiens high affinity choline transporter (SLC5A7) mRNA.
ACCESSION AF276871
VERSION AF276871.1 GI:10998441
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 1743)
Apparundaram,S., Ferguson,S.M., George,A.L., Jr. and Blakely,R.D.
Molecular cloning of a human, hemicholinium-3-sensitive choline
transporter
Biochem. Biophys. Res. Commun. 276 (3), 862-867 (2000)
JOURNAL 11027560
PUBMED 2 (bases 1 to 1743)
REFERENCE Apparundaram,S., Ferguson,S.M. and Blakely,R.D.
AUTHORS Direct Submission
TITLE Submitted (09-JUN-2000) Department of Pharmacology and Center for
JOURNAL Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at
Pierce, Nashville, TN 37232-6420, USA
FEATURES
source 1..1743
/organism="Homo sapiens"

```

gene
CDS
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q12"
1..1743
/gene="SLC5A7"
1..1743
/gene="SLC5A7"
/note="hCMT: solute carrier family 5 member 7"
/codon_start=1
/product="high affinity choline transporter"
/protein_id="AAG25940.1"
/db_xref="GI:10998442"
/translation="MAFVIEGLAIIVFLLILVGMAMRTKNGSAERSEALIV
GGRDGLVGFMTPTQVYKGRMSGLTIPLMGEMFMAAIFSLKGTISVITVD
AKPMRSKYVMTLDPQOYKGRMSGLTIPLMGEMFMAAIFSLKGTISVITVD
MHSVILSLALITLVVGLYSVATDVOLFCITVGLISVFPALSHPAVDIGFT
AVHAKYQKPMITVDSEVSWLDFLLMLGIPQAFORFLSSSSATVQVLSFL
AAGCLVMAIPAILIGALGASTDMNQAYLDPKTBESADMLIPVLQYLPVYSF
FELGASAVMSADSSILSASSMEANLYOSFRONASDKETVWWRITVFEFGSA
TAMALLTKTVGLAMYSPLVYIVIFPOLCVLPKGTNTVAGVAGVGLFRTIG
EYVLIQPLIFRYGYPDNGIYNQKPKITLAMYSPITNICTSLATLRESGLP
PFLVDVAVARSHENMDKTLIVKNENIKDLBALVKERQSHSTSTTNKAFILVD
DSSPDSGTEDNLQ"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 1743
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
Gaps: 0

US-10-724-806-4 (1-580) x AF276871 (1-1743)

QY 1 MetProPheHieValGluGlyLeuValAlaAlaIleLeuPheThyLeuLeuIlePheLeu 20
DB 1 ATGGCTTTCATGTGAGAGCATGATGATCATGTTCTCACTTCAATTTTGGCTG 60
QY 21 ValGlyIleThrAlaAlaTrpIlySerIlySaenSerGlyAsnProGluGluArgSerGlu 40
DB 61 GTTGGAAATATGGGCTCCCTGGAGAACCAAAACAGTGGCAGCCAGAAAGCCACGAA 120
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
DB 121 GCCATCATGTTGGTGGCCGAGATATTTGTTATTTGGTGGATTTACCAATGACAGCT 180
QY 61 ThrTrpValGlyGlyGlyIlyLeaenGlyThrAlaGluAlaValTyrglyProGlyCys 80
DB 181 ACCTGGGTGGAGAGGATATCATGACACAGCTGMAAGCAGTTTATGACAGGTTAT 240
QY 81 GlyLeuAlaIleThrAlaHisAlaProIleGlyTyrglySerLeuSerLeuIleLeuGlyGlyLeu 100
DB 241 GGCCTAGCTTGGGCTCAGGACCAACCAATTTGATTTCTTAGTCTGATTTTACGGTGGCTG 300
QY 101 PhePheAlaIlyProPheArgSerIlySerIlyValThrMetLeuSpproPheLeu 120
DB 301 TTCTTTGCAAAACCTATGCGTTCAAGGGGTATGTACCAATGTTACACCCGTTTCAGCAA 360
QY 121 IleTyrglyIlySaenSerGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
DB 361 ACTTATGAGAAACGATGGGCGGACCTCGTTTATTTCTGCACTGATGGAGAAATGTTTC 420
QY 141 ThrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleLeuAspVal 160
DB 421 TGGGCTGAGCAATTTTCTCTGCTTGGAGCCACCATCAGCGTGATCATCGATGGTGGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrglyThrLeuValGlyGly 180
DB 481 ATGCAATTTCTCTCATCTCTGCACTATGACATGTCACACTGTGACACCTGGTGGAGGG 540
QY 181 LeuTyrglySerValAlaTyrglyThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200

DB 541 CTCCTATTTCTGGCTTACACTGATGCTGACCTCTTTGATTTTGTAGGGCTGTGG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB 601 ATCAGGCTCCCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 221 HisAlaIlyTyrglySerProIlePheGlyIlyThrIleGlySerValGluValTyrglyTrp 240
DB 661 CATGCCAATATCCAAAGCCGCTGGGAGCACTGTTGATCTATGCAAGTCTTACTTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrglyPheGlnArg 260
DB 721 CTGTATAGTTTCTGTTGTTGATGCTGGGTGGAATCCATGCAACATATCTTTCAGAG 780
QY 261 ValLeuSerSerSerSerAlaThrTyrglyAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 781 GTTCTCTCTCTCTCTCCAGCCACTATGCTCAAGTGTCTCTCTCTGACGCTTTCGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
DB 841 TGCCTGTATATGCGATCCACCATCATCTATGCGGCACTTGGACATCAACAGCTGG 900
QY 301 AsnGlnThrAlaTyrglyTyrglyProAspProIlySerIlyGluGluAlaAspMetIleLeu 320
DB 901 AACCAAGCTGATATGGCTTCCAGATCCCAAGACTACAGAGAGGACAGATGATTTTA 960
QY 321 ProIleValLeuGlnTyrglyTyrglyCysProValTyrglySerPhePheIlyLeuGlyAlaVal 340
DB 961 CCATATGTTCTCAGATATCTGCTGCTGATATTTCTTCTGCTCTGCTGATGAGTT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1021 TCTGCTGCTGTTATGATCATCAGCAGATCTTTCATCTTGTACAGATTCATGCTTGA 1080
QY 361 ArgAsnIleTyrglyGlnLeuSerPheArgGlnAsnAlaSerAspIlyGluIleValThrVal 380
DB 1081 CGGAACATCATCAGCTTCTTCCAGCAAAATGCTTGGCAAAAGAAATCGTTGGGTT 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIly 400
DB 1141 ATGCGATACAGATGTTTGTGTTGGAGATCTGCAACAGCCATGCTTGTGCAAGAA 1200
QY 401 ThrValTyrglyLeuTrpTyrglyLeuSerSerAspLeuValTyrglyIleIlePheProGln 420
DB 1201 ACTGTATATGGGCTCTGATCTCAGTCTGATCTGATCTGATCTGATCTGATCTGATCTG 1260
QY 421 LeuLeuCysValLeuPheIleIlyGlyIlyThrAsnThrTyrglyAlaValAlaGlyTyrgly 440
DB 1261 CTGCTTGTGATCTTCTTGTAAAGGAAACCAACACTATGGGCGCTGGCAGGTTATGTT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrglyLeuTyrglyLeuGlnProLeuIle 460
DB 1321 TCTGGCTCTTCTGATGATTAACCTGAGGGAGCCATATCTGTATCTTACGCCCTTGATC 1380
QY 461 PheTyrglyTyrglyTyrglySerAspIlyTyrglyAsnGlnIlyTyrglyPheProPheIly 480
DB 1381 TTCTACCTGCTGATTTACCTGATGATGATGATGATGATGATGATGATGATGATGATG 1440
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrglyLeuAlaIlyTyrgly 500
DB 1441 ACACCTTGCATGTTATCATCTTCTTAAACCAACATTTGATCTCCATATCAGCAAGTAT 1500
QY 501 LeuPheGlySerGlyThrLeuProProIlySerLeuAspValPheAspAlaValAlaIlyArg 520
DB 1501 CTATTTGAAAGTGAACCTTGGCACTTGAATTAAGATTAAGATTAAGATTAAGATTAAG 1560
QY 521 HisSerGlyGluGlnMetAspIlyThrIleLeuValArgAsnGlnIlyAsnIleIlyLeuAsn 540
DB 1561 CACAGTGAAGAAACATGATTAAGCAATTTCTTGTCAAAATGCAAAATTAATTAAGAT 1620
QY 541 GluLeuAlaProValIlyProArgIlySerLeuThrLeuSerSerThrPheThrAsnIly 560

Db 1621 GAAGTTCAGCTTGACGACGACAGACGATGACCTCAGCTCAACTTACCAATATA 1680
Qy 561 GluAlaLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
Db 1661 GAGGCTTCCTTGATGATGATTCAGTCCAGAGAGGCTGTGGACATGAGATTAATTACAG 1740
RESULT 15
LOCUS HSA401466 1813 bp mRNA linear PRI 16-AUG-2000
DEFINITION Homo sapiens mRNA for high affinity choline transporter (CHTL gene).
ACCESSION AJ401466
VERSION AJ401466.1 GI:9843753
KEYWORDS CHTL gene; high affinity choline transporter.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Wieland, A., Bonisch, H. and Bruns, M.
TITLE Molecular cloning of the human and murine high affinity choline transporters and characterization of the human gene-structure
REFERENCE 2 (bases 1 to 1813)
AUTHORS Bruns, M.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2000) Bruns M., University of Bonn, Pharmacology and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY
FEATURES
source
1. 1813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q11-13"
/tissue_type="hypothalamus"
1. 1813
/gene="CHTL1"
19. 11761
/function="sodium- and chloride-dependent reuptake of choline"
/experiment="experimental evidence, no additional details recorded"
/codon_start=1
/product="high affinity choline transporter"
/protein_id="CAC03717.1"
/db_xref="GI:9843754"
/db_xref="GDB:096ZV3"
/translation="MAFHEGLIALIVLYLLLLGVIAAMETKNSSGAERSEAIIVGGRIGLVGGFTMTATVWGGYINGTAEAAYVPGYGLAMQAIPGYSLILGLFPAKPMKSKYVTMLDFPOIYGRMGLEFIPALMGEMWAAIFALGATTSIVLIDVNHISVITLALATYTLVGLSVATVVOVFCIFGLMISVPLFAGVAVGADIGFTAVHAKYQKPMICTYDSSSEVSWLDSFLLMGGITPMQAYFORVLSSSATYAOUSPLAAGCLVNAIPAILIGATGASTDMQNTAIGLPDKTIBEADILPTVLQYDQVYISFGALAVSAAMVSSADSSILSASMPARNIYQLSFRQMSDEKIVVMRIYVFGASATAMALLTFTVYGLWLTSSDLVYVIFPOLCVLPKGTNYGAVAGVSGFLRTGGEPYVLPGLIFPGYYPDDNGIYNOKPFPKTLAMYSFLTNICSYLAKYLESGTLPKLDVFDLVAVARHSEEMDKTILVKNENIKDELALVAVPRSMTLSTFTNKEATLDVDSPEGSGTEIDNLQ"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 1813
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
Gaps: 0
US-10-724-806-4 (1-580) x HSA401466 (1-1813)

Qy 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 19 ATGGCTTCCATGTGGAAGAGACTGATGATCATCGGTTCTCACTTAATTTCCTG 78
Qy 21 ValGlyIleTPrAlaAlaTPrIleThrIleAsnSerGlyAsnProGluGlnAspSerGlu 40
Db 79 GTTGAATATAGGCTGCTCGAGGAAACCAAAACAGTGGACGACGAGAGCGACGAA 138
Qy 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
Db 139 GCACATCATAGTTGGTGGCCGAGATATGCTTATTTGGTTGGATTAACATGACACT 198
Qy 61 ThrTrpValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db 199 ACCGTGGTCGAGGAGGCTATATCATGTCACAGCTGAAGCATTTATGTCACAGTTAT 258
Qy 81 GlyLeuAlaTPrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
Db 259 GGCTTAGCTGGCTCAGGACCAATTGGATATCTCTAGCTGATTTAGTGCTG 318
Qy 101 PhePheAlaIleProMetArgSerLeuGlyTyrValThrMetLeuAspProPheLeuGln 120
Db 319 TTCTTTGCAAAACCTATCGCTTCAAGGGATATGACCATGTTAGACCCGTTTCAGCA 378
Qy 121 IleTyrGlyIleArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140
Db 379 ATCTATGAGAAAACGATGCGGAGCTCGTTTATCTCTGACATGATGGAGAAATGTTTC 438
Qy 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 439 TGGCTGACGAAATTTCTCTGCTTTGGAGCCACATCAACGCGATATATGATGATGAT 498
Qy 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
Db 499 ATGCACATTTCTGTCATCATCTCGACATCTGCACTGTCATGTCACACTGGTGGAGGG 558
Qy 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db 559 CTCTATTCGTGGCTCATCAGTATGTCGTGAGCTCTTTCATTTTGTGGGGCTGG 618
Qy 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db 619 ATCAGCGTCCCTTTGCAATGTCATCTCGACATCGCAGACATGGGCTTCACTGCTGTG 678
Qy 221 HisAlaIleTyrGlnSerProTrpLeuGlyThrIleGlnSerValGluValTyrThrTrp 240
Db 679 CATGCCAATATCCAAACCCGCTGGGAGACTGTTGACTCATCTGAGTCTACTCTGG 738
Qy 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrPheGlnArg 260
Db 739 CTGTATAGTTTCTCTGTTTATAGTGTGGTGAATCCCATGGCAAGCAATCTTCAGAGG 798
Qy 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaAlaPheGly 280
Db 799 GTTCTCTCTTCTCTCTCGACCACTATGCTCAAGTGTCTCTCTCTGCGAGCTTTCGGG 858
Qy 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db 859 TGCTGGTGAATGCGCATCCAGCATACTCATTTGGGGCCATTTGGAGCATCAACAGACTGG 918
Qy 301 AenGlnThrAlaIleTyrGlyTyrProAspProIleThrIleGlyGluAlaAspMetLeuLeu 320
Db 919 AACCAAGCTGATATAGGGCTTCCAGATCCCAAGACTACAGAGAGGCGACATGATTTTA 978
Qy 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db 979 CCAATGTCTTCGACATATCTTCCCTGTGATATTTTCTTCTTGTGCTTGGTGGTCACTT 1038
Qy 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1039 TCTGCTGCTGTTATATCATCAGCAGATTTCTTCATCTTGTGACGAAATGTCATGTTTCA 1098

```
QY      361 ArgAsn11eTyrGlnuSerPheArgGlnAsn1aSerAspLysGlu11eVal1TrpVal 380
      |||
Db      1099 CGGAACATCTACACAGCTTTCCTTCAGACAAATCGTTGCGACAAAGAAATCGTTGGGTT 1158
QY      381 MetArg11eThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuThrLys 400
      |||
Db      1159 ATCCGATCAGACAGTGTGGTGGAGCATCTGCAACAGCCATGGCCTTGCTGACGAAA 1218
QY      401 ThrValTyrGlyLeuTrpLysSerSerAspLeuValTyr11e11e11ePheProGln 420
      |||
Db      1219 ACTGTATAGGGCTCTGGTACCTCAGTTCGACCTGTGTACATCGTATCTTCCCCAG 1278
QY      421 LeuLeuCyValLeuPhe11eLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyr11e 440
      |||
Db      1279 CTGCTTGTGTACTCTTGTGTAAAGGACCAACACATAGGGCCGTGGCAGGTATGTT 1338
QY      441 PheGlyLeuPheLeuArg11eThrGlyGlyProTyrLeuTyrLeuGlnProLeu11e 460
      |||
Db      1339 TCTGGCCTCTCTCGAATAACTGAGGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1398
QY      461 PheTyrProGlyTyrTyrSerAspLysAsnGly11eTyrAsnGlnArgPheProPheLys 480
      |||
Db      1399 TTCTACCTGGCTATTACCTGATGATATGATATATATATACAGAAATTCATTAA 1458
QY      481 ThrLeuSerMetValThrSerPhePheThrAsn11eCyValSerTyrLeuAlaLysTyr 500
      |||
Db      1459 ACACCTGCCATGCTTACATATTCTTAACCAACATTTGCATCTCTATCTAGCCAAAT 1518
QY      501 LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520
      |||
Db      1519 CTATTGAAGTGAACCTTGCCACCTAAATTAGATGATTTGATGCTGTTGTCAGA 1578
QY      521 HisSerGlnGluAsnMetAspLysThr11eLeuValArgAsnGluAsn11eLysLeuAsn 540
      |||
Db      1579 CACAGTGAAGAAACATGATTAAGACATCTGTCAAAATGAAATATTAATTAATGAT 1638
QY      541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
      |||
Db      1639 GAACCTTGCACTTGTGAAGCCACGACAGACATGACCTCGACTCACTTCAACCAATA 1698
QY      561 GluAlaLeuLeuAspValAspSerSerProGlnGlySerGlyThrGluAsnLeuGln 580
      |||
Db      1699 GAGGCTTCCTTGATGTGATCCAGTCCAGAAAGGCTGTGGACTGAAATTAATTAACAG 1758
```

Search completed: July 11, 2006, 02:03:20
Job time : 9566 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:25 ; Search time 1047 Seconds
(without alignments)
11607.096 Million cell updates/sec

Title: US-10-724-806-3

Perfect score: 1743

Sequence: 1 atgccttcctcatgctggaag.....ctgaagataatctacaatga 1743

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: N_Geneseq_8:*
2: Geneseq1980s:*
3: Geneseq1990s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*
14: Geneseq2005s:*
15: Geneseq2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1743	100.0	1743	10	ADD50640	ADD50640 CDNA enco
2	1743	100.0	1743	10	ADD50660	ADD50660 CDNA enco
3	1743	100.0	1743	14	ADV77898	ADV77898 Mouse hig
4	1743	100.0	1743	14	ADV77918	ADV77918 Mouse hig
5	1730.2	99.3	1743	4	AAFB1713	AAFB1713 Mouse hig
6	1730.2	99.3	1743	4	AAFB1713	AAFB1713 Mouse hig
7	1616.6	92.7	1743	4	AAFB1713	AAFB1713 Mouse hig
8	1616.6	92.7	1743	4	AAFB1713	AAFB1713 Mouse hig
9	1616.6	92.7	1743	4	AAFB1713	AAFB1713 Mouse hig
10	1378.2	79.1	1743	9	ABX94338	ABX94338 Human CDN
11	1375	78.9	1743	5	AAFA9207	AAFA9207 Human CHO
12	1375	78.9	1743	5	AAFA9207	AAFA9207 Human CHO
13	1375	78.9	1743	10	ADD50638	ADD50638 CDNA enco
14	1375	78.9	1743	14	ADV77896	ADV77896 Human hig
15	1375	78.9	1743	14	ADV77896	ADV77896 Human hig
16	1375	78.9	1743	14	ADV77896	ADV77896 Human hig
17	502.8	28.8	8760	5	AAHA9206	AAHA9206 Human CHO
18	502.8	28.8	26323	14	ABEB1646	ABEB1646 Human sci

19	502.8	28.8	119040	10	ADD50656	ADD50656 BAC seque
20	502.8	28.8	119040	14	ADV77914	ADV77914 Human BAC
21	502.8	28.8	142299	10	ADD50651	ADD50651 BAC seque
22	502.8	28.8	142299	14	ADV77909	ADV77909 Human BAC
23	388.8	22.3	1833	12	ADU93831	ADU93831 Human enco
24	356.4	20.4	1361	5	AAFD02461	AAFD02461 CDNA enco
25	355.8	20.4	1731	4	AAFB1710	AAFB1710 C. elegans
26	355.8	20.4	1985	10	ADD50644	ADD50644 High-affi
27	355.8	20.4	1985	14	ADV77902	ADV77902 Nematode
28	321	18.4	329	12	ADP08298	ADP08298 Subacute
29	278.8	16.0	1729	4	ABL29569	ABL29569 Drosophila
30	237.6	13.6	1461	6	ABK73210	ABK73210 Bacillus
31	178	10.2	4223	4	ABL29568	ABL29568 Drosophila
32	145.6	8.4	10140	5	AAHA9201	AAHA9201 Human CHO
33	145.6	8.4	119040	10	ADD50656	ADD50656 BAC seque
34	145.6	8.4	119040	14	ADV77914	ADV77914 Human BAC
35	145.6	8.4	142299	10	ADD50651	ADD50651 BAC seque
36	145.6	8.4	142299	14	ADV77909	ADV77909 Human BAC
37	132.4	7.6	455	4	AAI11913	AAI11913 Probe #18
38	132.4	7.6	455	4	ABA53620	ABA53620 Human fce
39	132.4	7.6	455	4	AAI33245	AAI33245 Probe #19
40	132.4	7.6	455	4	ABA23372	ABA23372 Probe #18
41	132.4	7.6	455	4	AAK27341	AAK27341 Human bon
42	132.4	7.6	455	4	AAK01883	AAK01883 Human bon
43	132.4	7.6	455	4	ABS26913	ABS26913 Human liv
44	132.4	7.6	455	5	AAI01851	AAI01851 Probe #18
45	132.4	7.6	1094	5	AA582193	AA582193 DNA encod

ALIGNMENTS

RESULT 1	
ADD50640	
ID	ADD50640 standard; CDNA; 1743 BP.
XX	
XX	ADD50640;
AC	
XX	
DT	15-JUN-2004 (first entry)
XX	
DE	CDNA encoding mouse high-affinity choline transporter (mCHT) #1.
XX	
XX	Mouse; high-affinity choline transporter; mCHT; cholinergic function;
KW	Parkinson's disease; Huntington's disease; Alzheimer's disease;
KW	schizophrenia; dysautonomia; myasthenia gravis; brain;
KW	cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;
KW	neuroprotective; neuroleptic; gene; ss.
XX	
OS	Mus sp.
XX	
FT	
FT	Key
FT	CDS
XX	
XX	
PN	US2003114399-A1.
XX	
PD	19-JUN-2003.
XX	
PF	23-JUL-2001; 2001US-00911077.
XX	
PR	23-JUL-2001; 2001US-00911077.
XX	
PA	(BLAK/) BLAKELY R D.
PA	(APPA/) APPARUNDARAM S.
PA	(FERG/) FERGUSON S.
XX	
XX	Blakely RD, Apparsundaram S, Ferguson S,
PI	
XX	
DR	WPI; 2003-810914/76.
XX	
XX	P-FSDB; ADD50641.
XX	
PT	Novel isolated polynucleotide encoding human or mouse high affinity

PT choline transporter polypeptide, useful in gene therapy to increase
PT cholinergic function in a cell of a patient suffering from Alzheimer's
PT disease.

XX Claim 30; SEQ ID NO 3; 74pp; English.

CC The present invention relates to the isolation of polynucleotide
CC sequences encoding human and mouse high-affinity choline transporter
CC (hCHT and mCHT respectively), and the proteins they encode. The gene
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
CC encoding hCHT is useful for expressing hCHT recombinantly. The hCHT
CC polynucleotide sequence when delivered to a cell, increases cholinergic
CC function in the cell that is in a patient having Parkinson's disease,
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
CC myasthenia gravis. The hCHT antibody is useful for controlling
CC transporter CHT proteins to the brain, and for treating the above
CC mentioned diseases. The antibody is also useful for diagnosing the above
CC mentioned disorders and to detect the influence of cholinergic
CC signaling. The present sequence encodes mCHT. Note: The sequence data
CC for this patent was obtained in electronic format directly from the USPTO
CC web site at seqdata.uspto.gov.

XX Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 10; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTTCCATGTTGGAAGAGCTGGTAGCTATTATCTTCTACCTCCTTATTTCTG 60
DB 1 ATGCTTTCCATGTTGGAAGAGCTGGTAGCTATTATCTTCTACCTCCTTATTTCTG 60
QY 61 GTTGAATATGAGGCTGCATGGAACCAAAACAGCGGCACCCAGAGAGCGCAGTGA 120
DB 61 GTTGAATATGAGGCTGCATGGAACCAAAACAGCGGCACCCAGAGAGCGCAGTGA 120
QY 121 GCCATCATAGTCGGGGGCGGTGACATGTTGTTGTTGTTGTTTACCATGACAGCC 180
DB 121 GCCATCATAGTCGGGGGCGGTGACATGTTGTTGTTGTTTACCATGACAGCC 180
QY 181 ACCTGGTGTGAGAGGCTACATCAATGAGCAGCAGAGAGCATGTATGGGCAAGTTGT 240
DB 181 ACCTGGTGTGAGAGGCTACATCAATGAGCAGCAGAGAGCATGTATGGGCAAGTTGT 240
QY 241 GGTCTAGCTTGGGCTCATGCAATGATGATTTCTCTGAGCTCTAATTTTAAAGTGTCTG 300
DB 241 GGTCTAGCTTGGGCTCATGCAATGATGATTTCTCTGAGCTCTAATTTTAAAGTGTCTG 300
QY 301 TTTTGTGGAACCTATGCGTTCGAAGGATATGATGATGTTGAACCATTCACAAACAG 360
DB 301 TTTTGTGGAACCTATGCGTTCGAAGGATATGATGATGTTGAACCATTCACAAACAG 360
QY 361 ATCTATGGAAGCGCATGAGTGGGCTGCTCTTCACTCCCTGCACTGATGGAGAGATGTTT 420
DB 361 ATCTATGGAAGCGCATGAGTGGGCTGCTCTTCACTCCCTGCACTGATGGAGAGATGTTT 420
QY 421 TGGGCTGAGCAATTTCTCTGCAATTAAGGGGCAACCATCAGCGTATCATATGTGAT 480
DB 421 TGGGCTGAGCAATTTCTCTGCAATTAAGGGGCAACCATCAGCGTATCATATGTGAT 480
QY 481 GTGAACATATGCGTATGTTCTCTGCACTTGGCATTCTTTAATACCTTAAGTGGTGGG 540
DB 481 GTGAACATATGCGTATGTTCTCTGCACTTGGCATTCTTTAATACCTTAAGTGGTGGG 540
QY 541 CTCTACTCTGTTGCAATATGATGATGTTGTCAGCTATTCCTGATTTTATAGAGCTGTGG 600
DB 541 CTCTACTCTGTTGCAATATGATGATGTTGTCAGCTATTCCTGATTTTATAGAGCTGTGG 600
QY 601 ATCAATGTCCTTTTGGCCCTGTCACATCTCTGAGTACACCGAATGGATTCACAGCTGTG 660
DB 601 ATCAATGTCCTTTTGGCCCTGTCACATCTCTGAGTACACCGAATGGATTCACAGCTGTG 660
QY 661 CATGCTAATATACCAAGTCCCTGCTGGGAGCAATGGAATCAGTTGAAGTTCACCTGG 720

DB 661 CATGCTAATATACCAAGTCCCTGCTGGGAGCAATGGAATCAGTTGAAGTTCACCTGG 720
QY 721 CTGATATATTTTCTGTTATATGATGCTGGGTGGAATCCCATGGCAACCTTCTCCAGAG 780
DB 721 CTGATATATTTTCTGTTATATGATGCTGGGTGGAATCCCATGGCAACCTTCTCCAGAG 780
QY 781 GTCTCTCTTATCTCTAGGCCAATGATGCTAGGTACTGTCTTCTCTGGCAGCTTTGGG 840
DB 781 GTCTCTCTTATCTCTAGGCCAATGATGCTAGGTACTGTCTTCTCTGGCAGCTTTGGG 840
QY 841 TGGCTGGATGAGGCTCTACCCGCCATATGATGATGAGAGCTATTGGAGTTCACAGACTGG 900
DB 841 TGGCTGGATGAGGCTCTACCCGCCATATGATGATGAGAGCTATTGGAGTTCACAGACTGG 900
QY 901 AACCAAGTCTCTACCGGTTATCCAGATCCAAAGACTAAGAGAGAGCAGACATGATCTC 960
DB 901 AACCAAGTCTCTACCGGTTATCCAGATCCAAAGACTAAGAGAGAGCAGACATGATCTC 960
QY 961 CCGATCTTCTGCAATGATCTCTGCTGTGTATCATCTCTCTTCTTGGGCTTGGTCTGT 1020
DB 961 CCGATCTTCTGCAATGATCTCTGCTGTGTATCATCTCTCTTCTTGGGCTTGGTCTGT 1020
QY 1021 TCAAGCTGCTCATGTCCTCTCAGCTGACTCGTCACATCTCTGCGGAGTTCTATGTTTCT 1080
DB 1021 TCAAGCTGCTCATGTCCTCTCAGCTGACTCGTCACATCTCTGCGGAGTTCTATGTTTCT 1080
QY 1081 CGGAATATCTACAGCTTCTCTTCCCTGACAAATGATCATGACAAAGAAATGTGTGGTCT 1140
DB 1081 CGGAATATCTACAGCTTCTCTTCCCTGACAAATGATCATGACAAAGAAATGTGTGGTCT 1140
QY 1141 ATGAGATCATCTGTCTTGTGTGGAGCATCTGCAACAGCCATGCTTGTCTGACGAAG 1200
DB 1141 ATGAGATCATCTGTCTTGTGTGGAGCATCTGCAACAGCCATGCTTGTCTGACGAAG 1200
QY 1201 ACTGTGTATGAGGCTCTGTGATCTGATCTGACCTGTGCTACATCATCATCTTCCACAG 1260
DB 1201 ACTGTGTATGAGGCTCTGTGATCTGATCTGACCTGTGCTACATCATCATCTTCCACAG 1260
QY 1261 CTGCTCTGTGTACTCTTCTCATCAAGAAACCAACCTTATGSGGCGAGTTGCTGTTATTT 1320
DB 1261 CTGCTCTGTGTACTCTTCTCATCAAGAAACCAACCTTATGSGGCGAGTTGCTGTTATTT 1320
QY 1321 TTTGAGCATATCTCGAGAAATTAAGAGAGAGCATCTATCTGACAGCCCTTATC 1380
DB 1321 TTTGAGCATATCTCGAGAAATTAAGAGAGAGCATCTATCTGACAGCCCTTATC 1380
QY 1381 TTCTAACCTGTTATTTACTCTGACAAAGATGATATCAATCAAGAGTTCACATTTTAA 1440
DB 1381 TTCTAACCTGTTATTTACTCTGACAAAGATGATATCAATCAAGAGTTCACATTTTAA 1440
QY 1441 ACTCTCTCATAGTGTACTCTCATTTCTTACCAACATTTGTTCTTATCTAGCCAAAGAT 1500
DB 1441 ACTCTCTCATAGTGTACTCTCATTTCTTACCAACATTTGTTCTTATCTAGCCAAAGAT 1500
QY 1501 CTATTTGAAAGTGAACCTTGCCTCCAAATTAAGTGAATGATGAGCTGTGTGCCAAGG 1560
DB 1501 CTATTTGAAAGTGAACCTTGCCTCCAAATTAAGTGAATGATGAGCTGTGTGCCAAGG 1560
QY 1561 CACAGTGAAGAGACATGAGACAAACATTTCTAGTCAGAAATGAAATATCAATTTAAAT 1620
DB 1561 CACAGTGAAGAGACATGAGACAAACATTTCTAGTCAGAAATGAAATATCAATTTAAAT 1620
QY 1621 GAACCTGCACTCTGTGAACCTTGGCAGAGCCTTAACTCAGTTCAACTTTCACCAATTAAG 1680
DB 1621 GAACCTGCACTCTGTGAACCTTGGCAGAGCCTTAACTCAGTTCAACTTTCACCAATTAAG 1680
QY 1681 GAGGCGCTCTGATGTTGATTTCCAGTCCGAGGAGGCTGTGGAGCTGAAGATTAATTTACA 1740
DB 1681 GAGGCGCTCTGATGTTGATTTCCAGTCCGAGGAGGCTGTGGAGCTGAAGATTAATTTACA 1740
QY 1741 TGA 1743

Db 1081 CGGAATATCTACCAAGCTTCTCTCAGACAAAATGCAATGAGCAAGAAATGTGTGGCTC 1140
QY 1141 ATGAGATACATGCTGCTGTGTGTTCCGAGCATCTGCAAGACCCATGGCTTTGGCTGACGAG 1200
Db 1141 ATGAGATACATGCTGCTGTGTGTTGGAGCATCTGCAAGACCCATGGCTTTGGCTGACGAG 1200
QY 1201 ACTGTGTATGGGCTCTGCTGCTGAGCTCTGACCTTGTCTACATCATCTTCCCAAG 1260
Db 1201 ACTGTGTATGGGCTCTGCTGAGCTCTGACCTTGTCTACATCATCTTCCCAAG 1260
QY 1261 CTGCTCTGTACTCTTCAATCAAGAACCAACTTATGGGGAGTGTCTGTATAT 1320
Db 1261 CTGCTCTGTACTCTTCAATCAAGAACCAACTTATGGGGAGTGTCTGTATAT 1320
QY 1321 TTGGACATATCTCTGAGAAATTAAGTGAAGAGCCATATCTTACTTGAAGCCCTTAATC 1380
Db 1321 TTGGACATATCTCTGAGAAATTAAGTGAAGAGCCATATCTTACTTGAAGCCCTTAATC 1380
QY 1381 TTCTACCTGTGTATTAATCTGACAGAAATGTATATACATCAAGAGTTCCTATTTAA 1440
Db 1381 TTCTACCTGTGTATTAATCTGACAGAAATGTATATACATCAAGAGTTCCTATTTAA 1440
QY 1441 ACTCTCTCCATGCTTACCTCATCTTCTTACCAACATTTGTGTTCTTATCTAGCCAAATAT 1500
Db 1441 ACTCTCTCCATGCTTACCTCATCTTCTTACCAACATTTGTGTTCTTATCTAGCCAAATAT 1500
QY 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATATGATATTTGATGCTGTGTGCGCAAG 1560
Db 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATATGATATTTGATGCTGTGTGCGCAAG 1560
QY 1561 CACGTGAAGAGAAATGAGACAGACCACTTCAAGTCAAGAAAGAAATATCAATTAAT 1620
Db 1561 CACGTGAAGAGAAATGAGACAGACCACTTCAAGTCAAGAAAGAAATATCAATTAAT 1620
QY 1621 GAACCTTGACCTGTGAAACCTGCGAGAGCCCTAACCTCAGTTCACTTACCATAAG 1680
Db 1621 GAACCTTGACCTGTGAAACCTGCGAGAGCCCTAACCTCAGTTCACTTACCATAAG 1680
QY 1681 GAGGCTCTCTTGAATGTTGATTCAGTCCGAGAGGGGTCTGGAGTGAAGATATTTACAA 1740
Db 1681 GAGGCTCTCTTGAATGTTGATTCAGTCCGAGAGGGGTCTGGAGTGAAGATATTTACAA 1740
QY 1741 TGA 1743
Db 1741 TGA 1743

RESULT 3
ADV77898
ID ADV77898 standard; cDNA, 1743 BP.
XX
AC ADV77898;
XX
DT 24-FEB-2005 (first entry)
XX
DE Mouse high affinity choline transporter (mCHT) cDNA.
XX
KW Choline transporter; neuromuscular disorder;
KW autonomic nervous system disorder; central nervous system disorder;
KW Parkinsons disease; Huntingtons chorea; genetic disorder;
KW Alzheimers disease; degeneration; neurological disease; schizophrenia;
KW psychiatric disease; myasthenia gravis; immune disorder; gene therapy;
KW cns-gen.; nocotropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW muscular-gen.; neuroleptic; DNA purification; gene; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT 1.1743
FT CDS /*tag= a
FT /product= "Mouse CHT protein"
XX
XX US2004248838-A1.

XX 09-DEC-2004.
PD 01-DEC-2003; 2003US-00724806.
XX 23-JUL-2001; 2001US-00911077.
XX (UYVA-) UNIV VANDERBILT.
PA Blakely RD, Apparsundaram S, Ferguson S;
PI MPI: 2005-020586/02.
XX P-PSDB; ADV77899.
XX DR GENBANK; AF276872.
XX
XX New human high affinity choline transporter (CHT) cDNA, useful for
PT treating a neuromuscular, autonomic or central nervous system disorder,
PT including Parkinson's disease, Huntington's disease, Alzheimer's,
PT schizophrenia.
PS Claim 30; SEQ ID NO 3; 73bp; English.
XX
XX The present invention provides polynucleotides encoding novel high
CC affinity choline transporters (CHTs), methods for their use in screening
CC and therapy. The invention is useful for treating neuromuscular
CC disorders, autonomic or central nervous system disorders such as
CC Parkinsons disease, Huntingtons chorea, Alzheimers disease,
CC schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is
CC also useful in gene therapy. The present sequence is mouse high affinity
CC choline transporter (mCHT) cDNA. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?docID=US20040248838.
XX
SQ Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;
Query Match 100.0%; Score 1743; DB 14; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AMGCCCTTCCATGCTGAGAGAGCTGGTAGCTATTAATCTCTTCACTCTTATATTTCTG 60
Db 1 ATGCCCTTCCATGCTGAGAGAGCTGGTAGCTATTAATCTCTTCACTCTTATATTTCTG 60
QY 61 GTTGAATATGCGGCTGATGGAAGAAACCAAAACAGCGGCAACCCAGAGCGCAGTGA 120
Db 61 GTTGAATATGCGGCTGATGGAAGAAACCAAAACAGCGGCAACCCAGAGCGCAGTGA 120
QY 121 GCCATCATATGTCGGGGGCGGTGACATTTGTTGTTGTTGTTTACATGACAGCC 180
Db 121 GCCATCATATGTCGGGGGCGGTGACATTTGTTGTTGTTTACATGACAGCC 180
QY 181 ACCTGGTGGAGAGAGGCTACATCATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACCTGGTGGAGAGAGGCTACATCATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 GGTCTAGCTTGGGCTCATGACACCAATTTGATTTCTCTGAGTCTATTTTATTTGTTG 300
Db 241 GGTCTAGCTTGGGCTCATGACACCAATTTGATTTCTCTGAGTCTATTTTATTTGTTG 300
QY 301 TTTTGTGGAAGACATATGCTTCAAGGAGATATGAGTATGAGTATGAGTATGAGTATGAG 360
Db 301 TTTTGTGGAAGACATATGCTTCAAGGAGATATGAGTATGAGTATGAGTATGAGTATGAG 360
QY 361 ATCTATGGAAG 420
Db 361 ATCTATGGAAG 420
QY 421 TGGGCTGAGCAATTTTCTCTGCAATTTAGGGGCAACATGATGATGATGATGAT 480
Db 421 TGGGCTGAGCAATTTTCTCTGCAATTTAGGGGCAACATGATGATGATGATGATGAT 480
QY 481 GTGAACATATGCGGCTATTTCTGCACTATGCAATTTCTTTATACCTTATGTTGGTGG 540

```

Db      ||||| 481 GGAACATATCGGTCATTCCTCTGCACTCATTCCTTAATCCCTAGGSGTGG 540
Qy      ||||| 541 CTTCTACTCTGTGGCATATCTGATGTTGTCCAGCTATTCGCACTTTTATAGACCTGTGG 600
Db      ||||| 541 CTTCTACTCTGTGGCATATCTGATGTTGTCCAGCTATTCGCACTTTTATAGACCTGTGG 600
Qy      ||||| 601 ATCAGAGTCCCTTTTCCCGTGCATCTCCGAGTACCGGACATCGGATTCACAGCTGTG 660
Db      ||||| 601 ATCAGAGTCCCTTTTCCCGTGCATCTCCGAGTACCGGACATCGGATTCACAGCTGTG 660
Qy      ||||| 661 CATGCTAAATACAGAGTCCCTGCTGGAGAACATTGAATCAAGTTGAAGTCTACACTGTG 720
Db      ||||| 661 CATGCTAAATACAGAGTCCCTGCTGGAGAACATTGAATCAAGTTGAAGTCTACACTGTG 720
Qy      ||||| 721 CTTGATTAATTTTCTGTTATTTGATGCTGGGTGGAAATCCCATGGCAAGCTTCCAGAGG 780
Db      ||||| 721 CTTGATTAATTTTCTGTTATTTGATGCTGGGTGGAAATCCCATGGCAAGCTTCCAGAGG 780
Qy      ||||| 781 GTCCCTCTTCATCCCTCAGCACTATGCTCAGAGTACTGTCCTTCCGAGAGCTTTGGG 840
Db      ||||| 781 GTCCCTCTTCATCCCTCAGCACTATGCTCAGAGTACTGTCCTTCCGAGAGCTTTGGG 840
Qy      ||||| 841 TGCCTGTGATGAGCTCTACCCGCTATATGATGAGCTATTTGAGCTTCCACAGACTGG 900
Db      ||||| 841 TGCCTGTGATGAGCTCTACCCGCTATATGATGAGCTATTTGAGCTTCCACAGACTGG 900
Qy      ||||| 901 AACCAAGCTGCTCAGCGGATATCCAGATCCCAAGACTTAAGAGGAGACAGACATGATTC 960
Db      ||||| 901 AACCAAGCTGCTCAGCGGATATCCAGATCCCAAGACTTAAGAGGAGACAGACATGATTC 960
Qy      ||||| 961 CGGATGCTCTGAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db      ||||| 961 CGGATGCTCTGAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy      ||||| 1021 TCAGCTGCTGATATGCTCTCAGCTGATCTGCTCCTGCTGCTGCTGCTGCTGCTGCT 1080
Db      ||||| 1021 TCAGCTGCTGATATGCTCTCAGCTGATCTGCTCCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy      ||||| 1081 CGGAATATCTACCAAGCTTCTTCAAGCAAAATGCAATCAGACAGAAATTTGTGGGTC 1140
Db      ||||| 1081 CGGAATATCTACCAAGCTTCTTCAAGCAAAATGCAATCAGACAGAAATTTGTGGGTC 1140
Qy      ||||| 1141 ATGAGGATCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db      ||||| 1141 ATGAGGATCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy      ||||| 1201 ACTGTGTATGAGCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1260
Db      ||||| 1201 ACTGTGTATGAGCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1260
Qy      ||||| 1261 CTGCTCTGTGATCTCTTCAATCAAGAGAACCACTTATGGGGCAGTGTGCTTATAT 1320
Db      ||||| 1261 CTGCTCTGTGATCTCTTCAATCAAGAGAACCACTTATGGGGCAGTGTGCTTATAT 1320
Qy      ||||| 1321 TTGGAATATCTGGAATATCTGGAAGAGAGCAATCTTAATCTTGAGAGCCCTTAATC 1380
Db      ||||| 1321 TTGGAATATCTGGAATATCTGGAAGAGAGCAATCTTAATCTTGAGAGCCCTTAATC 1380
Qy      ||||| 1381 TTCTACCTGTTATTTACTCTGACAAAGATGTATATACATAGAGGTTCCCATTTAA 1440
Db      ||||| 1381 TTCTACCTGTTATTTACTCTGACAAAGATGTATATACATAGAGGTTCCCATTTAA 1440
Qy      ||||| 1441 ACTCTCTCATAGGTTACCTCATCTTCTTACCAATTTGCTTCTTATCTAGCAAGTAT 1500
Db      ||||| 1441 ACTCTCTCATAGGTTACCTCATCTTCTTACCAATTTGCTTCTTATCTAGCAAGTAT 1500
Qy      ||||| 1501 CTAATTTGAAGTGAAGCTTGCCTCCAAATTTAGATGATTTGATGCTGTGCGAAGG 1560
Db      ||||| 1501 CTAATTTGAAGTGAAGCTTGCCTCCAAATTTAGATGATTTGATGCTGTGCGAAGG 1560
Qy      ||||| 1561 CACAGTGAAGAGAACTGACAAAGACCATTTCTGTGCAAGAAATGAAATATCAATTAAT 1620

```

```

Db      1561 CACAGTGAAGAGAACTGACAAAGACCATTTAGTCAGAAATGAAATATCAATTAAT 1620
Qy      1621 GAACCTTGACCTGTGAAACCTGGGAGAGGCTTAACCTTCAGTTCACTTCAACATAAG 1680
Db      1621 GAACCTTGACCTGTGAAACCTGGGAGAGGCTTAACCTTCAGTTCACTTCAACATAAG 1680
Qy      1681 GAGGCCCTCTTGATGTTGATTCAGTCCGAGAGGGGTCTGGAGCTGAGATAATTACAA 1740
Db      1681 GAGGCCCTCTTGATGTTGATTCAGTCCGAGAGGGGTCTGGAGCTGAGATAATTACAA 1740
Qy      1741 TGA 1743
Db      1741 TGA 1743

RESULT 4
ADV77918
ID ADV77918 standard; cDNA; 1743 BP.
XX
AC ADV77918;
XX
DT 24-FEB-2005 (first entry)
XX
DE Mouse high affinity choline transporter (mCHT) cDNA.
XX
KW Choline transporter; neuromuscular disorder;
KW autonomic nervous system disorder; central nervous system disorder;
KW parkinsons disease; huntingtons chorea; genetic disorder;
KW alzheimers disease; degeneration; neurological disease; schizophrenia;
KW psychiatric disease; myasthenia gravis; immune disorder; gene therapy;
KW cns-gen.; nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW muscular-gen.; neuroleptic; DNA purification; gene; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..1743
FT FT /*tag= a
FT FT /product= "Mouse CHT protein"
XX
PN US2004248838-A1.
XX
PD 09-DEC-2004.
XX
PF 01-DEC-2003; 2003US-00724806.
XX
PR 23-JUL-2001; 2001US-00911077.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Blakely RD, Apparsundaram S, Ferguson S;
XX
DR MPI; 2005-020586/02.
XX
DR P-PSDB; ADV77919.
XX
PT New human high affinity choline transporter (CHT) cDNA. useful for
PT treating a neuromuscular, autonomic or central nervous system disorder,
PT including Parkinson's disease, Huntington's disease, Alzheimer's,
PT schizophrenia.
XX
Example 4; SEQ ID NO 23; 73bp; English.
XX
The present invention provides polynucleotides encoding novel high
XX affinity choline transporters (CHTs), methods for their use in screening
XX and therapy. The invention is useful for treating neuromuscular
XX disorders, autonomic or central nervous system disorders such as
XX parkinsons disease, huntingtons disease, alzheimers disease,
XX schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is
XX also useful in gene therapy. The present sequence is mouse high affinity
XX choline transporter (mCHT) cDNA. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from the US patent office at
XX seqdata.uspto.gov/sequence.html?docid=US20040248838.

```

XX Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;
Query Match 100.0%; Score 1743; DB 14; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCTTTCCATGTGGAAGAGCTGGTATGATTTATCTCTTCTTACCTCTTATTTCTTG 60
1 ATGCTTTCCATGTGGAAGAGCTGGTATGATTTATCTCTTCTTACCTCTTATTTCTTG 60
1 ATGCTTTCCATGTGGAAGAGCTGGTATGATTTATCTCTTCTTACCTCTTATTTCTTG 60
61 GTTGAATATGGGCTGCATGTGAAAAACAAAACAGCGGCAACCAAGAGCGCAGTGA 120
61 GTTGAATATGGGCTGCATGTGAAAAACAAAACAGCGGCAACCAAGAGCGCAGTGA 120
61 GTTGAATATGGGCTGCATGTGAAAAACAAAACAGCGGCAACCAAGAGCGCAGTGA 120
121 GCCATCATAGTCGGGGGCGGTGACATTTGGTTTGTGTGTGTGTTTACATGACAGCC 180
121 GCCATCATAGTCGGGGGCGGTGACATTTGGTTTGTGTGTGTGTTTACATGACAGCC 180
121 GCCATCATAGTCGGGGGCGGTGACATTTGGTTTGTGTGTGTGTTTACATGACAGCC 180
181 ACCTGGGTTTGAAGAGCTACATCAATGGGACAGAGAGCAGTGTATGGGCGAGTTGT 240
181 ACCTGGGTTTGAAGAGCTACATCAATGGGACAGAGAGCAGTGTATGGGCGAGTTGT 240
181 ACCTGGGTTTGAAGAGCTACATCAATGGGACAGAGAGCAGTGTATGGGCGAGTTGT 240
241 GGTCTAGCTGGGCTCATGCAACCATTTGATTTCTGAGTCTAATTTTATAGGTTCTG 300
241 GGTCTAGCTGGGCTCATGCAACCATTTGATTTCTGAGTCTAATTTTATAGGTTCTG 300
241 GGTCTAGCTGGGCTCATGCAACCATTTGATTTCTGAGTCTAATTTTATAGGTTCTG 300
301 TTTTGTGGGAAACCTATGCGTTTCAAGGGATATGTACTATGTAGACCATTTCAACAG 360
301 TTTTGTGGGAAACCTATGCGTTTCAAGGGATATGTACTATGTAGACCATTTCAACAG 360
301 TTTTGTGGGAAACCTATGCGTTTCAAGGGATATGTACTATGTAGACCATTTCAACAG 360
361 ATCTATGGAAGCGCATGGGTGGGCTGCTTTCATCTCTGCACTGATGGGAGAGATGTT 420
361 ATCTATGGAAGCGCATGGGTGGGCTGCTTTCATCTCTGCACTGATGGGAGAGATGTT 420
361 ATCTATGGAAGCGCATGGGTGGGCTGCTTTCATCTCTGCACTGATGGGAGAGATGTT 420
421 TGGGCTGAGCAATTTCTCTGCAATTTAGGGGCGCACATCAGGGTATCATGATGAT 480
421 TGGGCTGAGCAATTTCTCTGCAATTTAGGGGCGCACATCAGGGTATCATGATGAT 480
421 TGGGCTGAGCAATTTCTCTGCAATTTAGGGGCGCACATCAGGGTATCATGATGAT 480
481 GTGAACATATGGGCTATGTTCTCTGCACTCATTTGCCATTTCTTATACCTTACGTTGG 540
481 GTGAACATATGGGCTATGTTCTCTGCACTCATTTGCCATTTCTTATACCTTACGTTGG 540
481 GTGAACATATGGGCTATGTTCTCTGCACTCATTTGCCATTTCTTATACCTTACGTTGG 540
481 GTGAACATATGGGCTATGTTCTCTGCACTCATTTGCCATTTCTTATACCTTACGTTGG 540
541 CTCTACTCTGTGGCATATCTGATGTGTCCAGCTATTTCTGCAATTTTATAGACTGTGG 600
541 CTCTACTCTGTGGCATATCTGATGTGTCCAGCTATTTCTGCAATTTTATAGACTGTGG 600
541 CTCTACTCTGTGGCATATCTGATGTGTCCAGCTATTTCTGCAATTTTATAGACTGTGG 600
601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGTGAGTCACCGCATGCGATTCACAGCTGT 660
601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGTGAGTCACCGCATGCGATTCACAGCTGT 660
601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGTGAGTCACCGCATGCGATTCACAGCTGT 660
661 CATGCTAATAATACAGAGTCCCTGTGGGAGAACCATTTGAATCAAGTTGAAGTCAACCTGG 720
661 CATGCTAATAATACAGAGTCCCTGTGGGAGAACCATTTGAATCAAGTTGAAGTCAACCTGG 720
661 CATGCTAATAATACAGAGTCCCTGTGGGAGAACCATTTGAATCAAGTTGAAGTCAACCTGG 720
661 CATGCTAATAATACAGAGTCCCTGTGGGAGAACCATTTGAATCAAGTTGAAGTCAACCTGG 720
721 CTGATTAATTTTCTGTATATGATGTGGGTGGAATCCCATGGCAAGCTTCAAGAGG 780
721 CTGATTAATTTTCTGTATATGATGTGGGTGGAATCCCATGGCAAGCTTCAAGAGG 780
721 CTGATTAATTTTCTGTATATGATGTGGGTGGAATCCCATGGCAAGCTTCAAGAGG 780
721 CTGATTAATTTTCTGTATATGATGTGGGTGGAATCCCATGGCAAGCTTCAAGAGG 780
781 GTCTCTCTTCAATCTCTGAGCCACCTATGCTCAGGTAATCTCTCTCTGAGAGCTTTTGGG 840
781 GTCTCTCTTCAATCTCTGAGCCACCTATGCTCAGGTAATCTCTCTCTGAGAGCTTTTGGG 840
781 GTCTCTCTTCAATCTCTGAGCCACCTATGCTCAGGTAATCTCTCTCTGAGAGCTTTTGGG 840
781 GTCTCTCTTCAATCTCTGAGCCACCTATGCTCAGGTAATCTCTCTCTGAGAGCTTTTGGG 840
841 TGCCCTGTGATGGCTCTACCCGCAATATGATAGAGGCTATTTGAGCTTCCACAGCTGG 900
841 TGCCCTGTGATGGCTCTACCCGCAATATGATAGAGGCTATTTGAGCTTCCACAGCTGG 900
841 TGCCCTGTGATGGCTCTACCCGCAATATGATAGAGGCTATTTGAGCTTCCACAGCTGG 900
841 TGCCCTGTGATGGCTCTACCCGCAATATGATAGAGGCTATTTGAGCTTCCACAGCTGG 900
901 AACCAAGACTGCTTACGGGATTCAGATCCCAAGACTTAAAGAGAGCAAGATGATTTCTC 960
901 AACCAAGACTGCTTACGGGATTCAGATCCCAAGACTTAAAGAGAGCAAGATGATTTCTC 960
901 AACCAAGACTGCTTACGGGATTCAGATCCCAAGACTTAAAGAGAGCAAGATGATTTCTC 960
901 AACCAAGACTGCTTACGGGATTCAGATCCCAAGACTTAAAGAGAGCAAGATGATTTCTC 960
961 CCGATGCTTCTGAGTACCTGCTGCTGTGATCATCTCTTCTTGGGCTTGTGCTGTT 1020
961 CCGATGCTTCTGAGTACCTGCTGCTGTGATCATCTCTTCTTGGGCTTGTGCTGTT 1020

Db 961 CCGATGCTTCTGAGTACCTGCTGCTGTGATCATCTCTTCTTGGGCTTGTGCTGTT 1020
Qy 1021 TCAGCTGTGATCATGCTCTGAGCTGACCTGTCATCTCTGTCGCGAGTTCTATGTTTCT 1080
Db 1021 TCAGCTGTGATCATGCTCTGAGCTGACCTGTCATCTCTGTCGCGAGTTCTATGTTTCT 1080
Qy 1081 CCGAATATCTACAGCTTCTCTTCAAGCAAAATGATAGAGCAAGAAATGTGTGGTCTC 1140
Db 1081 CCGAATATCTACAGCTTCTCTTCAAGCAAAATGATAGAGCAAGAAATGTGTGGTCTC 1140
Qy 1141 ATGAGGATACAGCTGCTGTTGTGTGAGCAATCTGCAACAGCCATGCTTGTGACGAG 1200
Db 1141 ATGAGGATACAGCTGCTGTTGTGTGAGCAATCTGCAACAGCCATGCTTGTGACGAG 1200
Qy 1201 ACTGTGATGAGGCTCTGCTGATCTGAGCTGTAACCTTGTCTATCATCATCTTCCACAG 1260
Db 1201 ACTGTGATGAGGCTCTGCTGATCTGAGCTGTAACCTTGTCTATCATCATCTTCCACAG 1260
Qy 1261 CTGCTGTGTGATCTCTTCAATGAAAGAACCAACCTTATGAGGCGAGTGTGTTATAT 1320
Db 1261 CTGCTGTGTGATCTCTTCAATGAAAGAACCAACCTTATGAGGCGAGTGTGTTATAT 1320
Qy 1321 TTTGAGCTATTCCTGAGAAATTAAGGAGAGGCAATCTATTAATCTGACGCCCTTAATC 1380
Db 1321 TTTGAGCTATTCCTGAGAAATTAAGGAGAGGCAATCTATTAATCTGACGCCCTTAATC 1380
Qy 1381 TTTCAACCTGCTTATTAATCTGACCAAGAGGATATATCAATCAGAGGTTCCCATTTTAA 1440
Db 1381 TTTCAACCTGCTTATTAATCTGACCAAGAGGATATATCAATCAGAGGTTCCCATTTTAA 1440
Qy 1441 ACTCTCTGATGATTAATCTGACCAAGAGGATATATCAATCAGAGGTTCCCATTTTAA 1440
Db 1441 ACTCTCTGATGATTAATCTGACCAAGAGGATATATCAATCAGAGGTTCCCATTTTAA 1440
Qy 1501 CTATTTGAAAGTGAACCTTGCCTCAAAATTAATGATGATGATGATGATGATGATGAT 1560
Db 1501 CTATTTGAAAGTGAACCTTGCCTCAAAATTAATGATGATGATGATGATGATGATGAT 1560
Qy 1561 CACAGTGAAGAGACATGAGCAAGACCAATCTAGTCAAGAAATGAAATATCAATTAAT 1620
Db 1561 CACAGTGAAGAGACATGAGCAAGACCAATCTAGTCAAGAAATGAAATATCAATTAAT 1620
Qy 1621 GAACCTGACCTGGAACCTGCGAGAGCCTTACCTCAGTCAACCTTACCAAAATTAAT 1680
Db 1621 GAACCTGACCTGGAACCTGCGAGAGCCTTACCTCAGTCAACCTTACCAAAATTAAT 1680
Qy 1681 GAGGCCCTCTTGAATGATTTCCAGTCCGAGGAGGCTGAGGACTGAAGATTAATTAACA 1740
Db 1681 GAGGCCCTCTTGAATGATTTCCAGTCCGAGGAGGCTGAGGACTGAAGATTAATTAACA 1740
Qy 1741 TGA 1743
Db 1741 TGA 1743

RESULT 5
AAFB1713
ID AAFB1713 standard; cDNA, 1743 BP.
XX AAFB1713;
AC 01-JUN-2001 (first entry)
XX
XX Mouse high affinity choline transporter protein encoding cDNA.
DE
XX
XX High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
KW ss.
XX
XX Mus musculus.
OS
FH Key Location/Qualifiers
FT 1..1743
CDS /*tag= a

FT /product= "high affinity choline transporter"
XX MO200116315-A1.
XX 08-MAR-2001.
XX 18-AUG-2000; 2000WO-JP005545.
XX 27-AUG-1999; 99JP-00240642.
XX 27-DEC-1999; 99JP-00368991.
XX (MISC-) JAPAN SCI & TECHNOLOGY CORP.
XX Haga T, Okuda T;
XX WPI: 2001-326688/23.
XX P-PSDB; AAB74666.
XX
XX New rat and human spinal cord high affinity choline transporters, useful
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
PT treating Alzheimer's disease.
XX
XX Claim 12; Page 78-82; 90pp; Japanese.
XX
XX The present sequence encodes a mouse (Mus musculus) high affinity choline
CC transporter protein designated cho-1. The cho-1 protein has nootropic and
CC neuroprotective activities. The cho-1 polynucleotide and protein can be
CC used for the diagnosis of diseases related to the expression of cho-1 by
CC comparing the cho-1 polynucleotide sequence in a sample to that of a
CC control. Drug compositions containing the cho-1 protein or expression
CC promoters or inhibitors of cho-1 are useful for treating disorders
CC characterised by abnormal levels of cho-1, such as Alzheimer's disease
XX
SQ Sequence 1743 BP; 407 A; 410 C; 409 G; 517 T; 0 U; 0 Other;

Query Match 99.3%; Score 1730.2; DB 4; Length 1743;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1735; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGCCCTTCCATGTTGGAAGAGCTGTGATATATCTCTTCACTCTTATATTTCTG 60
DB 1 ATGCTCTTCCAGTAGAAGAGCTGTGATATATCTCTTCACTCTTATATTTCTG 60
QY 61 GTTGGAAATATGGGCTCATGAAAAACCAAAAACAGGGCAACCCAGAAGCGCAGTGA 120
DB 61 GTTGGAAATATGGGCTCATGAAAAACCAAAAACAGGGCAACCCAGAAGCGCAGTGA 120
QY 121 GCCATCATAGTCGGGGCCGTGACATTTGTTGTTGGTTGTTTTCATGACAGCC 180
DB 121 GCCATCATAGTCGGGGCCGTGACATTTGTTGTTGGTTTTCATGACAGCC 180
QY 181 ACCCTGGTTGGAGAGGCTACATCAATGGGACAGCAAGCATGTATGGGCGAAGTTGT 240
DB 181 ACCCTGGTTGGAGAGGCTACATCAATGGGACAGCAAGCATGTATGGGCGAAGTTGT 240
QY 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTGTCTG 300
DB 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTGTCTG 300
QY 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTACTATGTATGACCATTTCAACAG 360
DB 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTACTATGTATGACCATTTCAACAG 360
QY 361 ATCTATGAAAGGAGATGGGTGGGCTCTTCATCCCTGCACTGATGGAGAGATGTC 420
DB 361 ATCTATGAAAGGAGATGGGTGGGCTCTTCATCCCTGCACTGATGGAGAGATGTC 420
QY 421 TGGGCTGACAGCAATTTCTCTGCAATTAAGGCGCACCATCAGCGTGATCAATTAAT 480
DB 421 TGGGCTGACAGCAATTTCTCTGCAATTAAGGCGCACCATCAGCGTGATCAATTAAT 480
QY 481 GTGAACATATCGGTCTTGTCTTGCACTCATTTGCACTTTTATACCTTATGGGTGGG 540
DB 481 GTGAACATATCGGTCTTGTCTTGCACTCATTTGCACTTTTATACCTTATGGGTGGG 540

DB 481 GTGAACATATCGGTCTTGTCTTGCACTCATTTGCACTTTTATACCTTATGGGTGGG 540
QY 541 CTCTACTCGTGGCAATATGATGTCGAGCTATCTGATTTTATAGAGCTGGG 600
DB 541 CTCTACTCGTGGCAATATGATGTCGAGCTATCTGATTTTATAGAGCTGGG 600
QY 601 ATCAGTGTCCCTTTTGGCCCTGTACATCCGTGAGTCAACGACATGCGATTCACGCTGTG 660
DB 601 ATCAGTGTCCCTTTTGGCCCTGTACATCCGTGAGTCAACGACATGCGATTCACGCTGTG 660
QY 661 CATGCTAAATACAGAGTCCCTGGGACCAATTTGAATCAAGTTGAAGTCAACCTGG 720
DB 661 CATGCTAAATACAGAGTCCCTGGGACCAATTTGAATCAAGTTGAAGTCAACCTGG 720
QY 721 CTTGATTAATTTTCTGTTATGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAGG 780
DB 721 CTTGATTAATTTTCTGTTATGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAGG 780
QY 781 GTCTCTCTTCACTCTCAAGCCACTATGCTCAGTATCTGCTCTTCTGACGCTTTTGGG 840
DB 781 GTCTCTCTTCACTCTCAAGCCACTATGCTCAGTATCTGCTCTTCTGACGCTTTTGGG 840
QY 841 TGCCTGTGATGGCTCTTACCCGCCATATGCAATAGAGCTATTGGACCTTCCACAGCTGG 900
DB 841 TGCCTGTGATGGCTCTTACCCGCCATATGCAATAGAGGCTATTGGAGCTTCCACAGCTGG 900
QY 901 AACCGACGCTCTAGGGGATCCAGATCCCAAGATCAAGAGAGGACAGATGATTC 960
DB 901 AACCGACGCTCTAGGGGATCCAGATCCCAAGATCAAGAGAGGACAGATGATTC 960
QY 961 CCGATGCTTCTGACAGTACCTGCTGCTGTGATCATCTCTTGGGCTGGTGTCT 1020
DB 961 CCGATGCTTCTGACAGTACCTGCTGCTGTGATCATCTCTTGGGCTGGTGTCT 1020
QY 1021 TTAGCTGCTGATGCTCTCACTGACCTGCTGCTCATCTGCTGCGAGCTTATGTTGCT 1080
DB 1021 TTAGCTGCTGATGCTCTCACTGACCTGCTGCTCATCTGCTGCGAGCTTATGTTGCT 1080
QY 1081 CGGAATATCTACAGCTTCCCTTCAAGCAAAATGCAATGCAAGAAATTTGTTGGGTC 1140
DB 1081 CGGAATATCTACAGCTTCCCTTCAAGCAAAATGCAATGCAAGAAATTTGTTGGGTC 1140
QY 1141 ATGAGGATCACTGTGCTGTGTTGTTGGAGCATCTGCAACGCCATGGCTTGTGCAAG 1200
DB 1141 ATGAGGATCACTGTGCTGTGTTGTTGGAGCATCTGCAACGCCATGGCTTGTGCAAG 1200
QY 1201 ACTGTGTATGGGCTCTGGTACTGAGCTTGACCTTGTCTAATCATCTTCCACAG 1260
DB 1201 ACTGTGTATGGGCTCTGGTACTGAGCTTGACCTTGTCTAATCATCTTCCACAG 1260
QY 1261 CTGCTCTGTGTACTCTTCAACAAAGGAAACCACTTATAGGGGCAAGTTGCTGTTATTT 1320
DB 1261 CTGCTCTGTGTACTCTTCAACAAAGGAAACCACTTATAGGGGCAAGTTGCTGTTATTT 1320
QY 1321 TTTGACATATCTGAGATTAATCTGAGAGAGGACCATCTATCTTATGAGCCCTTAATC 1380
DB 1321 TTTGACATATCTGAGATTAATCTGAGAGAGGACCATCTATCTTATGAGCCCTTAATC 1380
QY 1381 TTCTACCCCTGGTTATTAATCTGACAAAGATGATATACATCAAGAGTTCCCATTTAA 1440
DB 1381 TTCTACCCCTGGTTATTAATCTGACAAAGATGATATACATCAAGAGTTCCCATTTAA 1440
QY 1441 ACTCTCTCCAGGTTAATCCCTTATTAACCAATTTGCTTATCTATAGCAAGTAT 1500
DB 1441 ACTCTCTCCAGGTTAATCCCTTATTAACCAATTTGCTTATCTATAGCAAGTAT 1500
QY 1501 CTATTTGAAAGTGAACCTTGGCTCCAAAATTAATGATATTTTATGCTGTGTGCAAG 1560
DB 1501 CTATTTGAAAGTGAACCTTGGCTCCAAAATTAATGATATTTTATGCTGTGTGCAAG 1560
QY 1561 CACAGTGAAGAGAAATGAGCAAGCAACCTTCAAGTCAAGAAATGAAATTTAAAT 1620
DB 1561 CACAGTGAAGAGAAATGAGCAAGCAACCTTCAAGTCAAGAAATGAAATTTAAAT 1620


```
QY 1621 GAACCTGACCTGTGAAAACCTCGGACAGCCTTAACCTCAGTTCACTTTCACCATTAAG 1680
    |||
Db 1621 GAACCTGACCTGTGAAAACCTCGGACAGCCTTAACCTCAGTTCACTTTCACCATTAAG 1680
QY 1681 GAGGCCCCCTGTTGATGATTCAGTCCGAGGAGGCTGTGGACCTGAAGATTAATTACAA 1740
    |||
Db 1681 GAGGCCCCCTGTTGATGATTCAGTCCGAGGAGGCTGTGGACCTGAAGATTAATTACAA 1740
QY 1741 TGA 1743
    |||
Db 1741 TGA 1743

RESULT 6
AAD02457
ID AAD02457 standard; cDNA; 4938 BP.
AC AAD02457;
XX
XX 24-APR-2001 (first entry)
DT
DE Mouse P4P6B1 OMA (obese mice adipocyte) protein encoding cDNA.
XX
XX Mouse; OMA protein; obese mice adipocyte; P4P6B1;
KW fuel metabolism disorder; therapy; obesity; diabetes; gene therapy;
KM anorectic; antidiabetic; ss.
XX
OS Mus sp.
FH Key Location/Qualifiers
FT CDS 247..1989
FT CDS /*tag= a
FT misc_feature /product= "Mouse P4P6B1 OMA protein"
FT /*tag= b
FT /*note= "Portion of original 450 bp PCR fragment"
XX
XX WO200078950-A2.
XX
XX 28-DEC-2000.
XX
XX 13-JUN-2000; 2000MO-US016217.
XX
XX 22-JUN-1999; 99US-0141515P.
XX
XX (AMYL-) AMYLIN PHARM INC.
PA
PI Sierzege M, Albrandt K;
PI WPI: 2001-112322/12.
DR P-PDB; AA72388.
XX
XX Novel obese mice adipocyte polypeptides useful in diagnosis and treatment
PT of disorders of fuel metabolism such as obesity or diabetes.
XX
XX Claim 2; Fig 3; 83pp; English.
XX
XX The present sequence is mouse P4P6B1 cDNA which encodes OMA (obese mice
CC adipocyte) protein. The P4P6B1 fragment was generated by RNA
CC fingerprinting using random primers P4 and P6. OMA is used as a
CC diagnostic reagent for diagnosing a disorder of fuel metabolism in an
CC underweight or an overweight individual, by detecting the transcription
CC level of a gene encoding OMA, which is induced or repressed in the
CC individual by a factor such as genetic obesity, fasting and refeeding of
CC a fasted individual. OMA is useful in the generation of antibodies, for
CC use in pharmaceutical compositions and for studying DNA/protein
CC interactions. Nucleic acids encoding OMA are involved in gene therapy. An
CC inhibitor of OMA or an antisense oligonucleotide that inhibits expression
CC of OMA are useful for treating disorders of fuel metabolism such as
CC obesity or diabetes
XX
XX Sequence 4938 BP; 1436 A; 1012 C; 976 G; 1514 T; 0 U; 0 Other;
```

```
Query Match 99.3%; Score 1730.2; DB 5; Length 4938;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1735; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGCCCTTCCATGTCGGAAGACCTGTAGCTATATATCCCTTCTCACTTCCTATATTTCTG 60
    |||
Db 247 ATGCTTTCCACGTAGAGAGACCTGTAGCTATATATCCCTTCTCACTTCCTATATTTCTG 306
QY 61 GTTGAAATATGAGGCTGATGGAACCAACCAAAACAGCGCAACCCAGAAAGCGCAGTGA 120
    |||
Db 307 GTTGAAATATGAGGCTGATGGAACCAACCAAAACAGCGCAACCCAGAAAGCGCAGTGA 366
QY 121 GCCATCATAGTCGGGGGCGCGTGAATTTGTTGGTGGTGGTTTACATGACAGCC 180
    |||
Db 367 GCCATCATAGTCGGGGGCGCGTGAATTTGTTGGTGGTGGTTTACATGACAGCC 426
QY 181 ACCTGGGTTGAGAGAGGCTCATCATATGAGACAGAGAGAGATGATGGCCAGTTGT 240
    |||
Db 427 ACCTGGGTTGAGAGAGGCTCATCATATGAGACAGAGAGAGATGATGGCCAGTTGT 486
QY 241 GGTCTAGCTTGGGCTCATGACCCATTTGATATTTCTGACTCTAATTTTAAAGTGTCTG 300
    |||
Db 487 GGTCTAGCTTGGGCTCATGACCCATTTGATATTTCTGACTCTAATTTTAAAGTGTCTG 546
QY 301 TTTTGGGAAACCTATGCGTTCCAGAGGATATGACTATGTAAGACCCATTCAGAG 360
    |||
Db 547 TTTTGGGAAACCTATGCGTTCCAGAGGATATGACTATGTAAGACCCATTCAGAG 606
QY 361 ATCTATGGAAGAGCCATGAGTGGGCTGCTCTTCAATCCCTGACATGAGAGATGTTT 420
    |||
Db 607 ATCTATGGAAGAGCCATGAGTGGGCTGCTCTTCAATCCCTGACATGAGAGATGTTT 666
QY 421 TGGGCTGACGAATTTTCTGCAATGAGGGCCACATCAGCGTATATTGATGTGAT 480
    |||
Db 667 TGGGCTGACGAATTTTCTGCAATGAGGGCCACATCAGCGTATATTGATGTGAT 726
QY 481 GTGAACATATGGGTCATGCTCTGCACTCATTCGCAATCTTTTAACTGTGGGTGG 540
    |||
Db 727 GTGAACATATGGGTCATGCTCTGCACTCATTCGCAATCTTTTAACTGTGGGTGG 786
QY 541 CTCTACTCTGGAATATAGATGTTGTCAGGATATCTGACATTTTATAGACTGTGG 600
    |||
Db 787 CTCTACTCTGGAATATAGATGTTGTCAGGATATCTGACATTTTATAGACTGTGG 846
QY 601 ATCAGTGTCCCTTTTGGCTGTCAATCTGCAATCTGACATCTGACATCTGAGCTGTG 660
    |||
Db 847 ATCAGTGTCCCTTTTGGCTGTCAATCTGCAATCTGACATCTGAGCTGTG 906
QY 661 CATGCTAATATCAGAGTCCCTGGCTGGGAACATGTAATGATGTAAGTCAACCTGG 720
    |||
Db 907 CATGCTAATATCAGAGTCCCTGGCTGGGAACATGTAATGATGTAAGTCAACCTGG 966
QY 721 CTGTGTAATTTTCTGTATATGATGCTGGGTGTAATCCCATGGCAAGCCTACTTCAGAG 780
    |||
Db 967 CTGTGTAATTTTCTGTATATGATGCTGGGTGTAATCCCATGGCAAGCCTACTTCAGAG 1026
QY 781 GTCTCTCTTATCTCTACGCCCATTAATGCTCAGGTACTGCTCTTCTGTGGACCTTTTGGG 840
    |||
Db 1027 GTCTCTCTTATCTCTACGCCCATTAATGCTCAGGTACTGCTCTTCTGTGGACCTTTTGGG 1086
QY 841 TGCCGTGATGAGGCTCTACCGGCCCATATGATGATGAGAGTATTTGAGTCTTCACAGACTGG 900
    |||
Db 1087 TGCCGTGATGAGGCTCTACCGGCCCATATGATGATGAGAGTATTTGAGTCTTCACAGACTGG 1146
QY 901 AACCAAGACTGCTTACGGGTATTCAGATCCAAAGCTAAGAGGAAGACAGACATGTTCTC 960
    |||
Db 1147 AACCAAGACTGCTTACGGGTATTCAGATCCAAAGCTAAGAGGAAGACAGACATGTTCTC 1206
QY 961 CCGATCGTTTTCAGACTACTCTGCGCTGTGTACATCTCTTCTTGGGCTGTGTCTGT 1020
    |||
Db 1207 CCGATCGTTTTCAGACTACTCTGCGCTGTGTACATCTCTTCTTGGGCTGTGTCTGT 1266
```



```

OY 1021 TAGCTGCTGTCATGTCCTCAAGTGAAGCTGCTCATCTGTGCGGAGTTCTATGTTGCT 1080
DB 1267 TCAGCTGCTGTCATGTCCTCAAGTGAAGCTGCTCATCTGTGCGGAGTTCTATGTTGCT 1326
OY 1081 CGGAATATCTACAGCTTCTCTTCAAGAAATGCTCAAGAAATGCTGCTGCT 1140
DB 1327 CGGAATATCTACAGCTTCTCTTCAAGAAATGCTCAAGAAATGCTGCTGCT 1386
OY 1141 ATGAGGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1387 ATGAGGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
OY 1201 ACTGCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1447 ACTGCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1506
OY 1261 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 1507 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566
OY 1321 TTGGAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
DB 1567 TTGGAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1626
OY 1381 TTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB 1627 TTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
OY 1441 ACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 1687 ACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746
OY 1501 CTATTTGAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
DB 1747 CTATTTGAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1806
OY 1561 CACAGTGAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 1807 CACAGTGAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1866
OY 1621 GAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 1867 GAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1926
OY 1681 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 1927 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1986
OY 1741 TGA 1743
DB 1987 TGA 1989

```

RESULT 7
AAAF81711
ID AA81711 standard; cDNA; 1743 BP.

AAAF81711;

01-JUN-2001 (first entry)

Rat high affinity choline transporter protein encoding cDNA.

High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;

OS Rattus norvegicus.

Key Location/Qualifiers

FT 1.1743

FT /product= "high affinity choline transporter"

```

PN WO200116315-A1.
XX
PD 08-MAR-2001.
XX
PF 18-AUG-2000; 2000WO-JP005545.
XX
PR 27-AUG-1999; 99JP-00240642.
XX
PR 27-DEC-1999; 99JP-00368991.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Haga T, Okuda T;
XX
XX WPI; 2001-226688/23.
DR P-PSDB; AAB74664.
XX
PT New rat and human spinal cord high affinity choline transporters, useful
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
PT treating Alzheimer's disease.
XX
XX Claim 6; Page 64-68; 90pp; Japanese.
XX
CC The present sequence encodes a rat (Rattus norvegicus) high affinity
CC choline transporter protein designated cho-1. The cho-1 protein has
CC neurotrophic and neuroprotective activities. The cho-1 polynucleotide and
CC protein can be used for the diagnosis of diseases related to the
CC expression of cho-1 by comparing the cho-1 polynucleotide sequence in a
CC sample to that of a control. Drug compositions containing the cho-1
CC protein or expression promoters or inhibitors of cho-1 are useful for
CC treating disorders characterized by abnormal levels of cho-1, such as
CC Alzheimer's disease
XX
SQ Sequence 1743 BP; 414 A; 402 C; 404 G; 523 T; 0 U; 0 Other;

```

Query Match 92.7%; Score 1616.6; DB 4; Length 1743;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```

OY 1 ATGCTCTTCATGCTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
DB 1 ATGCTCTTCATGCTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
OY 61 GTTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 GTTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
OY 121 GGCATCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 121 GGCATCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
OY 181 ACCTGGGCTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 ACCTGGGCTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
OY 241 GGTCTAGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 GGTCTAGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
OY 301 TTTTTCGCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 TTTTTCGCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
OY 361 ATCTATGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 ATCTATGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
OY 421 TGGGCTGCAAGCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 TGGGCTGCAAGCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
OY 481 GTGAACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 GTGAACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

```

```

Qy 541 CTCTACTGTGGCATATATCTGATGTGTGTCAGCTATTCCTGACATTTTATATAGACGTGG 600
Db 541 CTCTACTGTGGCATATATCTGATGTGTGTCAGCTATTCCTGACATTTTATATAGACGTGG 600
Qy 601 ATGAGTGTCCCTTTGGCCCTGTGTCACATCCTGAGTGCACCCGACATGGGATTTCAGCTGG 660
Db 601 ATGAGTGTCCCTTTGGCCCTGTGTCACATCCTGAGTGCACCCGACATGGGATTTCAGCTGG 660
Qy 661 CATGCTAAATAACAGAGTCCCTGCTGGGAAACCATTTGATCAGTTGAGTCTACACTGG 720
Db 661 CATGCTAAATAACAGAGTCCCTGCTGGGAAACCATTTGATCAGTTGAGTCTACACTGG 720
Qy 721 CTGTAAATTTTCTGTATTGATGCTGGGTGAATCCATGGCAAGCTTCCAGAG 780
Db 721 CTGTAAATTTTCTGTATTGATGCTGGGTGAATCCATGGCAAGCTTCCAGAG 780
Qy 781 GTCCCTCTCTTCAATCCTGAGCCACCTATGCTCAGTAATCTCTTCTGCGACCTTTGGG 840
Db 781 GTCCCTCTCTTCAATCCTGAGCCACCTATGCTCAGTAATCTCTTCTGCGACCTTTGGG 840
Qy 841 TGCCCTGTGATGGCTCTACCCGCAATGCTATGAGAGCTTGGAGCTTCCACAGACTGG 900
Db 841 TGCCCTGTGATGGCTCTACCCGCAATGCTATGAGAGCTTGGAGCTTCCACAGACTGG 900
Qy 901 AACCAAGACTGCTGACGGGTATCCAGATCCCAAGACTTAAAGAGAGACAGATGATTC 960
Db 901 AACCAAGACTGCTGACGGGTATCCAGATCCCAAGACTTAAAGAGAGACAGATGATTC 960
Qy 961 CGGATGCTTGGACATCCTGCTGCTGTGATCATCTCTTGGGTGGGTGTGCTGT 1020
Db 961 CGGATGCTTGGACATCCTGCTGCTGTGATCATCTCTTGGGTGGGTGTGCTGT 1020
Qy 1021 TCAGCTGTGATGATGCTCTGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 TCAGCTGTGATGATGCTCTGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 1081 CGGAATATCTACAGCTTCTTCCAGCAAAATGCAATGCAAGAGAAATGTGTGGGTC 1140
Db 1081 CGGAATATCTACAGCTTCTTCCAGCAAAATGCAATGCAAGAGAAATGTGTGGGTC 1140
Qy 1141 ATGAGGATCAGTGTGCTGTGCTGAGATGCTGCAAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 ATGAGGATCAGTGTGCTGTGCTGAGATGCTGCAAGAGAGAGAGAGAGAGAGAGAG 1200
Qy 1201 ACTGTGTATGGGCTGTGATCTGAGCTGAGCTGCTGTGCTGATCATCATCTTCCACAG 1260
Db 1201 ACTGTGTATGGGCTGTGATCTGAGCTGAGCTGCTGTGCTGATCATCATCTTCCACAG 1260
Qy 1261 CTGCTGTGTGATCTTCTTCAAAAGAACCACTTATGGGGCAGTTCCTGTTATAT 1320
Db 1261 CTGCTGTGTGATCTTCTTCAAAAGAACCACTTATGGGGCAGTTCCTGTTATAT 1320
Qy 1321 TTGTGACTATTCCTGAGATTACTGGAGAGAGCCATATCTTACTTGTGAGCCCTTAATC 1380
Db 1321 TTGTGACTATTCCTGAGATTACTGGAGAGAGCCATATCTTACTTGTGAGCCCTTAATC 1380
Qy 1381 TTCTAACCTGTATTACTCTGACAAAGATGATATACATCAGAGGTTCCCATTTAA 1440
Db 1381 TTCTAACCTGTATTACTCTGACAAAGATGATATACATCAGAGGTTCCCATTTAA 1440
Qy 1441 ACTCTCTCATAGTATACCTCATCTTTTACCAACATTTGTGTTCTTATCTAGCCAGTAT 1500
Db 1441 ACTCTCTCATAGTATACCTCATCTTTTACCAACATTTGTGTTCTTATCTAGCCAGTAT 1500
Qy 1501 CTATTGAAAGTGAAGCTTGGCTCCAAATAGATGATTTGATGCTGTGGCAGAG 1560
Db 1501 CTATTGAAAGTGAAGCTTGGCTCCAAATAGATGATTTGATGCTGTGGCAGAG 1560
Qy 1561 CACAGTGAAGAGACATGAGCAAGACATTTAGTCAAGAAATGAAAATCAATTTAAAT 1620
Db 1561 CACAGTGAAGAGACATGAGCAAGACATTTAGTCAAGAAATGAAAATCAATTTAAAT 1620

```

```

Qy 1621 GAACCTTGACCTGTGAAAACCTGGGAGAGCCTTAACCTCAGTTCAACTTTCACAAATAAG 1680
Db 1621 GAACCTTGACCTGTGAAAACCTGGGAGAGCCTTAACCTCAGTTCAACTTTCACAAATAAA 1680
Qy 1681 GAGGCCCTCTTGATGATTTCCAGTCCGAGAGGGGCTGGGAGCTGAAGATTAATTACA 1740
Db 1681 GAGGCCCTCTTGATGATTTCCAGTCCGAGAGGGATCTGGGAGCTGAAGATTAATTACA 1740
Qy 1741 TGA 1743
Db 1741 TGA 1743

RESULT 8
ADD50642
ID ADD50642 standard; cDNA; 4904 BP.
XX
AC ADD50642;
XX
DT 15-JAN-2004 (first entry)
XX
DE cDNA encoding rat high-affinity choline transporter (rCHT).
XX
KW Rat, high-affinity choline transporter; rCHT; cholinergic function;
KW Parkinson's disease; Huntington's disease; Alzheimer's disease;
KW schizophrenia; dysautonomia; myasthenia gravis; brain;
KW cholinergic signaling; antiparkinsonian; anticonvulsant; nootropic;
KW neuroprotective; neuroleptic; gene; ss.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT CDS 224..1966
FT /tag= a
FT /product= "rCHT"
XX
PN US2003114399-A1.
XX
PD 19-JUN-2003.
XX
PF 23-JUL-2001; 2001US-00911077.
XX
PR 23-JUL-2001; 2001US-00911077.
XX
PA (BLAKELY R D.
PA (APPA/) APPARUNDARAM S.
PA (FERG/) FERGUSON S.
XX
PI Blakely RD, Apparsundaram S, Ferguson S;
XX
DR WPI; 2003-810914/76.
XX
DR P-PSDB; ADD50643.
XX
PT Novel isolated polynucleotide encoding human or mouse high affinity
PT choline transporter polypeptide, useful in gene therapy to increase
PT cholinergic function in a cell of a patient suffering from Alzheimer's
PT disease.
XX
PS Example 1; SEQ ID NO 5; 74pp; English.
XX
CC The present invention relates to the isolation of polynucleotide
CC sequences encoding human and mouse high-affinity choline transporter
CC (hCHT and mCHT respectively), and the proteins they encode. The gene
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
CC encoding hCHT is useful for expressing hCHT recombinantly. The hCHT
CC polynucleotide sequence when delivered to a cell, increases cholinergic
CC function in the cell that is in a patient having Parkinson's disease,
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
CC myasthenia gravis. The hCHT antibody is useful for controlling
CC transporter CHT proteins in the brain, and for treating the above
CC mentioned diseases. The antibody is also useful for diagnosing the above
CC mentioned disorders and to detect the influence of cholinergic
CC signalling. The present sequence encodes rat CHT (rCHT). Note: The

```

CC sequence data for this patent was obtained in electronic format directly
 from the USPTO web site at segdata.uspto.gov.

XX Sequence 4904 BP; 1447 A; 991 C; 939 G; 1527 T; 0 U; 0 Other;

Query Match 92.7%; Score 1616.6; DB 10; Length 4904;
 Best Local Similarity 95.5%; Pred. No. 0;
 Matches 1664; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```

QY 1 ATGCTTTTCATGTGGAAGAGCTGCTGCTATTTCTCTTCTACCTCTCTATTTCTG 60
DB 224 ATGCTTTTCATGTGGAAGAGCTGCTGCTATTTCTCTTCTACCTCTCTATTTCTG 283
QY 61 GTTGAATATGAGGCTGCTGGAAGAAACCAAAACAGGCGCAACCCAGAGAGGCTGAA 120
DB 284 GTTGAATATGAGGCTGCTGGAAGAAACCAAAACAGGCGTAATGCAAGAAACGACGAA 343
QY 121 GGCATCATAGTCGGGGGCGGTGACATTTGTTGTTGTTGTTTTCATGACAGCC 180
DB 344 GGCATCATAGTTGGGGGCGGAGCATTGTTGTTGTTGTTTTCATGACAGCC 403
QY 181 ACCGTGGTGGAGAGGCTACATCAATGGGACAGCAAGCACTGATAGGCGCAAGTTGT 240
DB 404 ACCGTGGTGGAGAGGCTACATCAACGCGGACAGCTGAAGCACTTATAGGCGCAAGTTGT 463
QY 241 GGTCTAGCTTGGGCTGATGCAACCATGATATTCCTGAGTCTAATTTTAAAGTGTCTG 300
DB 464 GGTCTAGCTTGGGCTGATGCAACCATGATATTCCTGAGTCTAATTTTAAAGTGTCTG 523
QY 301 TTTTTCGCAAACTATGCGTTCAGAGGATATGCTATGTTAGACCATTTCAAAACAG 360
DB 524 TTTTTCGCAAACTATGCGTTCAGAGGATATGCTATGTTAGACCATTTCAAAACAG 583
QY 361 ATCTATGAAAGCGCATGGGTGGCTGCTCTTCTATCCCTGCACTGATGGAGAGATGTT 420
DB 584 ATCTATGAAAGCGCATGGGTGGCTGCTCTTCTATCCCTGCACTGATGGAGAGATGTT 643
QY 421 TGGGGTGGCAAAATTTCTCTGCTATGAGGGGCGACCATGAGGCTGATGATGAGAT 480
DB 644 TGGGGTGGCAAAATTTCTCTGCTATGAGGGGCGCTACATGAGGCTGATGATGAGAT 703
QY 481 GTGAACATATCGGTCAATTTCTCTGCACTATTTCCATTTCTTATACCTTATGGGTGG 540
DB 704 GTGAACATATCGGTCAATTTCTCTGCACTATTTCCATTTCTTATACCTTATGGGTGG 763
QY 541 CTCTACTCTGTGGCATATGCTGATGTTGTCAGCTATTTCTGCACTTTTATAGAGCTGT 600
DB 764 CTCTACTCTGTGGCATATGCTGATGTTGTCAGCTATTTCTGCACTTTTATAGAGCTGT 823
QY 601 ATCAGTGTCCCTTTGGCCGTGCATCTCTGCAAGTACCGGACATGGGATTTCAAGCTGT 660
DB 824 ATCAGTGTCCCTTTGGCCGTGCATCTCTGCAAGTACCGGACATGGGATTTCAAGCTGT 883
QY 661 CATGCTAAATACAGAGTCCCTGGCTGGGAAACCATTTGAATCAGTTGAAGCTTACACCT 720
DB 884 CATGCTAAATACAGAGTCCCTGGCTGGGAAACCATTTGAATCAGTTGAAGCTTACACCT 943
QY 721 CTGTATATATTTCTGTATTTGATGCTGGGTGGAATCCCATGGACCTTACAGAG 780
DB 944 CTGTATATATTTCTGTATTTGATGCTGGGTGGAATCCCATGGACCTTACAGAG 1003
QY 781 GTTCCTCTCTCATCCCTGAGCACTATGCTCAGGTACTGCTTCTCTGGAGCTTTTGG 840
DB 1004 GTTCCTCTCTCATCCCTGAGCACTATGCTCAGGTACTGCTTCTCTGGAGCTTTTGG 1063
QY 841 TGGCTGTGATGCTCTACCCGCGCATATGATAGAGACTTATTTGAGCTTTCACAGACT 900
DB 1064 TGGCTGTGATGCTCTACCCGCGCATATGATAGAGACTTATTTGAGCTTTCACAGACT 1123
QY 901 AACCAACGCTCAGAGGATACAGATCCCAAGACTAAGAGAGAGAGAGAGATGATTTCT 960
DB 1124 AACCAACGCTCAGAGGATACAGATCCCAAGACTAAGAGAGAGAGAGATGATTTCT 1183
  
```

```

QY 961 CCGATCGTCTGACAGTACTCTGCCCCCTGTGTACATCTCTTCTTTGGGCTGTGCTGT 1020
DB 1184 CCGATGTTTCTACAGTACTCTGCCCCCTGTGTACATCTCTTCTTTGGGCTGTGCTGT 1243
QY 1021 TTAGCTGCTGATGATCTCTGACCTGATCTGCTCATCTCTGCGGAGTTCTATGTTGCT 1080
DB 1244 TTAGCTGCTGATGATCTCTGACCTGATCTGCTCATCTCTATCAGCAAGTTCAATGTTGCT 1303
QY 1081 CGGAATATCTACAGTCTCTCTGACCAAAATGATCAGACAGAGAAATTTGTGGCTC 1140
DB 1304 CGGAATATCTACAGTCTCTCTGACCAAAATGATCAGACAGAGAAATTTGTGGCTC 1363
QY 1141 ATGAGGATCACTGTCTGTGTGCGAGCATCTGCAACGCCATGCTTGTGTCAGAG 1200
DB 1364 ATGAGGATCACTGTGTGTGTGCGAGCATCTGCAACGCCATGCTTGTGTCAGAG 1423
QY 1201 ACTGTGTATGGGCTCTGCTGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1424 ACTGTGTATGGGCTCTGCTGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1483
QY 1261 CTGCTCTGTGTACTCTTCACTAAAGGAAACCACTTATGGGCGAGTTGCTGTTATTT 1320
DB 1484 CTGCTCTGTGTACTCTTCACTAAAGGAAACCACTTATGGGCGAGTTGCTGTTATTT 1543
QY 1321 TTTGACATATTCCTGAGATTTACTGAGAGAGAGCCATATCTTATCTTGACGCCCTTAT 1380
DB 1544 TTTGACATATTCCTGAGATTTACTGAGAGAGAGCCATATCTTATCTTGACGCCCTTAT 1603
QY 1381 TTCTACCTCTGTTATTTACTCTGACAAAGATGATATACATCAGAGTTCCCATTTAA 1440
DB 1604 TTCTACCTCTGTTATTTACTCTGACAAAGATGATATACATCAGAGTTCCCATTTAA 1663
QY 1441 ACTCTCTGATGATGATACCTGATTTCTTACCAATTTGTTCTTATGAGCAAGTAT 1500
DB 1664 ACTCTCTGATGATGATACCTGATTTCTTACCAATTTGTTCTTATGAGCAAGTAT 1723
QY 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTTAGATGTTTGTGATGCTGTGCGAAG 1560
DB 1724 CTATTTGAAAGTGAACCTTGGCTCCAAATTTAGATGTTTGTGATGCTGTGCTCAAG 1783
QY 1561 CACAGTAAAGAGAACATGAGCAAGACCATTTCTGATGAGAAATGAAATTTCAATTTAAT 1620
DB 1784 CACAGTAAAGAGAACATGAGCAAGACCATTTCTGATGAGAAATGAAATTTCAATTTAAT 1843
QY 1621 GAATCTGACCTGTGAAACCTGCGGAGAGCTTAACCTGATCACTTCACTCAATTAAG 1680
DB 1844 GAATCTGACCTGTGAAACCTGCGGAGAGCTTAACCTGATCACTTCACTCAATTAAG 1903
QY 1681 GAGGCCCTCTGATGTTGATTTCAAGTCCGAGAGGCTGTGGACTGAGATATTTTACAA 1740
DB 1904 GAGGCCCTCTGATGTTGATTTCAAGTCCGAGAGGCTGTGGACTGAGATATTTTACAA 1963
QY 1741 TGA 1743
DB 1964 TGA 1966
  
```

RESULT 9

ADV77900
 ID ADV77900 standard; cDNA; 4904 BP.

ADV77900;

24-FEB-2005 (first entry)

Rat high affinity choline transporter (rCHT1) cDNA.

Choline transporter; neuromuscular disorder;
 autonomic nervous system disorder; central nervous system disorder;
 Parkinson's disease; Huntington's chorea; genetic disorder;
 Alzheimer's disease; degeneration; neurological disease; schizophrenia;
 psychiatric disease; myasthenia gravis; immune disorder; gene therapy;
 CNS-gen.; nootropic; neuroprotective; antiparkinsonian; anticonvulsant;

QY 1441 ACTCTCCATGGTACCTCATCTTTACCAATTTGTTCTTATCTAGCAGTAT 1500
 DB 1664 ACTCTCCATGGTACCTCATCTTTACCAATTTGTTCTTATCTAGCAGTAT 1723
 QY 1501 CTAATTTGAAGTGAACCTTGGCTCCAAATTAATGATGATTTGATCTGTTGCGAAG 1560
 DB 1724 CTAATTTGAAGTGAACCTTGGCTCCAAATTAATGATGATTTGATCTGTTGCGAAG 1783
 QY 1561 CACAGGAGAGAACATGAGCAAGACCTTCTAGTAAATGAATTAATCAATTAAT 1620
 DB 1784 CACAGGAGAGAACATGAGCAAGACCTTCTAGTAAATGAATTAATCAATTAAT 1843
 QY 1621 GAATCTGACCTGTGAAACCTGCGAGAGCCTTAACCTGATTCATCTTACCAATAAG 1680
 DB 1844 GAATCTGACCTGTGAAACCTGCGAGAGCCTTAACCTGATTCATCTTACCAATAAG 1903
 QY 1681 GAGGCCCTCTTGATGTTGATTCAGTCCGAGAGGCTTGGAGCTGAAGTAATTTACAA 1740
 DB 1904 GAGGCCCTCTTGATGTTGATTCAGTCCGAGAGGCTTGGAGCTGAAGTAATTTACAA 1963
 QY 1741 TGA 1743
 DB 1964 TGA 1966
 RESULT 10
 ABX94338
 ID ABX94338 standard; cDNA, 1743 BP.
 AC ABX94338;
 XX
 DT 13-JUN-2003 (first entry)
 XX
 DE Human cDNA encoding high affinity choline transporter, HACT.
 XX
 KW Human; ss; gene; HACT; high affinity choline transporter; pain;
 KW neurotransmitter biosynthesis; learning and memory; aging; epilepsy;
 KW neurological disorder; spasticity; myoclonus; muscle spasm;
 KW muscle hyperactivity; stroke; head trauma; neuronal cell death;
 KW multiple sclerosis; spinal cord injury; dystonia; Alzheimer's disease;
 KW Myasthenia Gravis; multi-infarct dementia; AIDS dementia;
 KW Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KW ALS; attention deficit disorder; organic brain syndrome; schizophrenia;
 KW nicotine addiction; memory disorder; cognitive disorder.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1743
 FT /tag= a
 FT /product= "HACT"
 FT
 XX
 PN US6500643-B1.
 XX
 PD 31-DEC-2002.
 XX
 PF 07-SEP-2000; 2000US-00657252.
 XX
 PR 07-SEP-2000; 2000US-00657252.
 XX
 PA (UFL) UNIV FLORIDA.
 XX
 PI Wu D, Gu Y, Millard WJ, He Y;
 XX
 DR WPI, 2003-361355/34.
 XX
 DR P-PSDB; AB008979.
 XX
 PT Novel isolated polynucleotide (1) that encodes high affinity choline
 PT transporter protein, useful for preventing, treating or ameliorating
 PT neurological and cognitive disorders such as Alzheimer's or Parkinson's
 PT disease.
 XX
 PS Claim 2; Col 17-21; 20pp; English.

XX The invention relates to an isolated polynucleotide which encodes a high
 CC affinity choline transporter (HACT) protein appearing as AB008979. Also
 CC included are a polynucleotide encoding a fragment consisting of at least
 CC about 50 amino acids of the HACT protein, a vector comprising the
 CC polynucleotide which comprises at least about 12 contiguous nucleic acids
 CC of a polynucleotide appearing as ABX94338 (encoding choline
 CC acetyltransferase), a recombinant host cell which comprises the vector
 CC (used to express the HACT protein or fragment). The polynucleotide is
 CC useful as a probe or primer to detect the presence of HACT polynucleotide
 CC in a sample, such as a biological sample, or for screening for test
 CC agents which bind to the polynucleotide. A pharmaceutical composition
 CC comprising the polynucleotide is useful for preventing, treating or
 CC ameliorating neurological and cognitive disorders e.g. pain, spasticity,
 CC myoclonus, muscle spasm, muscle hyperactivity, epilepsy, stroke, head
 CC trauma, neuronal cell death, multiple sclerosis, spinal cord injury,
 CC dystonia, Alzheimer's disease, myasthenia gravis, multi-infarct
 CC dementia, AIDS dementia, Parkinson's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis (ALS), attention deficit disorder, nicotine
 CC addiction, organic brain syndrome, schizophrenia or memory and cognitive
 CC disorders. HACT is thought to be the rate limiting step in cholinergic
 CC neurotransmitter biosynthesis and regeneration (cholinergic transmissions
 CC are crucial to brain functions such as learning and memory). The present
 CC sequence encodes human HACT
 CC
 SO Sequence 1743 BP; 411 A; 395 C; 405 G; 532 T; 0 U; 0 Other;
 Query Match 79.1%; Score 1378.2; DB 9; Length 1743;
 Best Local Similarity 86.9%; Pred. No. 0;
 Matches 1515; Conservative 0; Mismatches 228; Indels 0; Gaps 0;
 QY 1 ATGCCCTTCCATGTGAGAGAGCTGTGATATATCTCTTACCTCCTTAATTTCTG 60
 DB 1 ATGCCCTTCCATGTGAGAGAGCTGTGATATATCTCTTACCTCCTTAATTTCTG 60
 QY 61 GTTGGAAATATGAGGCTGATGAGAAACCAAAACAGGCGAACCCAGAGAGCCAGTGA 120
 DB 61 GTTGGAAATATGAGGCTGATGAGAAACCAAAACAGGCGAACCCAGAGAGCCAGTGA 120
 QY 121 GCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 DB 121 GCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 181 ACCGTGGCTGAGAGAGAGCTGTGATATATCTCTTACCTCCTTAATTTCTG 240
 DB 181 ACCGTGGCTGAGAGAGAGCTGTGATATATCTCTTACCTCCTTAATTTCTG 240
 QY 241 GGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 DB 241 GGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 301 TTTTGTGCAAACTATGCGTTCCAAAGGATATGATGATGATGATGATGATGATGAT 360
 DB 301 TTTTGTGCAAACTATGCGTTCCAAAGGATATGATGATGATGATGATGATGATGAT 360
 QY 361 ATCTATGAGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 DB 361 ATCTATGAGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 421 TGGGCTGAGCAATTTTCTGATATGAGGCGCACCATGATGATGATGATGATGAT 480
 DB 421 TGGGCTGAGCAATTTTCTGATATGAGGCGCACCATGATGATGATGATGATGAT 480
 QY 481 GTGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 481 GTGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 541 CTCTACTGTGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 541 CTCTACTGTGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 601 ATCAGTGTCCCTTTGGCCCTGTCATCTGAGTACCGAGCATGCGGATTCACAGCTGTG 660


```
Db      121 GGCATCATAGTGGTGGCCGAGATATGGTTATTGGTTGGTGGATTTCCATGACAGCT 180
Qy      181 ACCGTGGTGGAGAGGCTACATCAATGGGACAGCAAGAGCTGATATGGCCAGTTGT 240
Db      181 ACCGTGGTGGAGAGGCTATATCAATGGCAGCAGCTGAAGCAGTTATATGATACAGTTAT 240
Qy      241 GGTCTAGCTTGGGCTCATGACCCCATTTGATATATCTGAGTCTAATTTTAGTGCTG 300
Db      241 GGCCTAGCTTGGGCTCAGGACCAATTTGATATCTTATGCTGATTTTAGTGCTG 300
Qy      301 TTTTTCGCAAACTATGCGCTTCAAGGATATGTAATGTTAGACCATTTCAAAACAG 360
Db      301 TTTTTCGCAAACTATGCGCTTCAAGGATATGTAATGTTAGACCATTTCAAAACAG 360
Qy      361 ATCTATGAAAAGCGCATGGGTGGGCTCTTCACTGCACTGATGGGAGAGATGTTG 420
Db      361 ATCTATGAAAAGCGCATGGGTGGGCTCTTCACTGCACTGATGGGAGAGATGTTG 420
Qy      421 TGGGCTGAGCAATTTTCTGCAATTAAGGGGCCACCATGAGGCTGATCAATTTGATG 480
Db      421 TGGGCTGAGCAATTTTCTGCTGTTGGAGCCACCATAGGCTGATCAATGATGAT 480
Qy      481 GTGAACATATCGGTCAATGCTCTGCACTATTTGCCATTTCTTATACCTAGTGAGT 540
Db      481 ATGCAATTTCTGTATCATCTCTGCACTATTTGCCATCTGTACACATGGTGGAGGG 540
Qy      541 CTCTACTCTGTGGCATATCTGATGTTGTCAGCTATTTCTGCAATTTTATAGACCTGTG 600
Db      541 CTCTACTCTGTGGCATATCTGATGTTGTCAGCTATTTCTGCAATTTTATAGACCTGTG 600
Qy      601 ATCAGTGTCCCTTTGGCCCTGTCACATCTGCACTGAGCAGCAGCATTTGCACTGAGT 660
Db      601 ATCAGTGTCCCTTTGGCCCTGTCACATCTGCACTGAGCAGCAGCATTTGCACTGAGT 660
Qy      661 CATGCTAAATACAGAGTCCCTGGCTGGGAGACCATTTGAATCACTTGAAGTCAACCTG 720
Db      661 CATGCTAAATACAGAGTCCCTGGCTGGGAGACCATTTGAATCACTTGAAGTCAACCTG 720
Qy      721 CTTGATATTTTCTGTTATGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAG 780
Db      721 CTTGATATTTTCTGTTATGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAG 780
Qy      781 GTCCCTCTTCATCCCTCAGCAGCTATGCTCAGGTATGCTGCTTCTGAGAGCTTTGG 840
Db      781 GTTCCTCTCTTCTCTCAGCAGCTATGCTCAGGTATGCTGCTTCTGAGAGCTTTGG 840
Qy      841 TGCCTGTGATGAGCTTACCCGCTATATGATAGAGCTATTTGAGCTTCCACAGACTG 900
Db      841 TGCCTGTGATGAGCTTACCCGCTATATGATAGAGCTATTTGAGCTTCCACAGACTG 900
Qy      901 AACCAAGACTGCTTACCGGCTATCCAGATCCCAAGACTTAAGAGAGAGCAGATGATTCT 960
Db      901 AACCAAGACTGCTTACCGGCTATCCAGATCCCAAGACTTAAGAGAGAGCAGATGATTCT 960
Qy      961 CCGATGCTCTGAGTACCTGCGCTGCTGATCATCTCTTGGGGCTGGTGTGCTGT 1020
Db      961 CCAATGCTCTGAGTATCTGCGCTGCTGATCATCTCTTGGGGCTGGTGTGCTGT 1020
Qy      1021 TCAGCTGTGCTATGCTCAGCTGATGCTGCTCATCTGCGGAGTTCTATGTTGCT 1080
Db      1021 TCAGCTGTGCTATGCTCAGCTGATGCTGCTCATCTGCGGAGTTCTATGTTGCT 1080
Qy      1081 CGGAATATCTACAGCTTCTTCAACAAGATGATGATGATGATGATGATGATGATG 1140
Db      1081 CGGAATATCTACAGCTTCTTCAACAAGATGATGATGATGATGATGATGATGATG 1140
Qy      1141 ATGAGATCACTGCTGCTGCTGCTGAGAGATGCAAGAGCAGCTTGTGAGCAAG 1200
Db      1141 ATGAGATCACTGCTGCTGCTGCTGAGAGATGCAAGAGCAGCTTGTGAGCAAG 1200
Qy      1201 ACTGTATAGGCTGTGATGCTGAGCTTGTGATCATCATCTTCCACAG 1260
Db      1201 ACTGTATAGGCTGTGATGCTGAGCTTGTGATCATCATCTTCCACAG 1260
```

```
Db      1201 ACTGTATAGGCTGTGATGCTGAGCTTGTGATCATCATCTTCCACAG 1260
Qy      1261 CTGCTCTGCTGCTCTCTTCAACAAGATGATGATGATGATGATGATGATGATG 1320
Db      1261 CTGCTCTGCTGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1320
Qy      1321 TTTGAGCTATCTCTGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db      1321 TTTGAGCTATCTCTGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy      1381 TTTTACCTGCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Db      1381 TTTTACCTGCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Qy      1441 ACTGCTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Db      1441 ACTGCTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Qy      1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTAATGATGATGATGATGATGATGATG 1560
Db      1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTAATGATGATGATGATGATGATGATG 1560
Qy      1561 CACAGTGAAGAGATGAGACATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db      1561 CACAGTGAAGAGATGAGACATGATGATGATGATGATGATGATGATGATGATGATG 1620
Qy      1621 GAAGCTGCACTGCTGAAACCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db      1621 GAAGCTGCACTGCTGAAACCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Qy      1681 GAGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
Db      1681 GAGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
Qy      1741 TGA 1743
Db      1741 TGA 1743

RESULT 12
AAH49207
ID AAH49207 standard; cDNA; 1743 BP.
XX
AC AAH49207;
XX
DT 26-NOV-2001 (first entry)
XX
DE Human CHOT encoding cDNA.
XX
KW CHOT; human; choline transporter; chromosome 2q11-13; nootropic;
KW neuroprotective; gene therapy; antisense therapy; degenerative disease;
KW cognitive disorder; Alzheimer's disease; ss.
XX
OS Homo sapiens.
XX
PN DE1000905-A1.
XX
PD 30-AUG-2001.
XX
PF 28-FEB-2000; 2000DE-01009055.
XX
PR 28-FEB-2000; 2000DE-01009055.
XX
PA (BRUE/) BRUESS M.
PA (BOEN/) BOENISCH H.
XX
PI Brues M., Boenisch H;
XX
XX WPI; 2001-590709/67.
XX DR P-PSDB; AAB6837.
XX
XX A new gene encoding human choline transporter, designated hCHOT is
XX located on chromosome 2q11-13 and is useful to treat degenerative
PT
```


PT disorders such as Alzheimer's disease.

XX Disclosure; Page 11; 12pp; German.

XX This invention describes a novel gene encoding human choline transporter, designated hCHOT which is located on chromosome 2q11-13. The products of the invention have nootropic and neuroprotective activity and can be used for gene or antisense therapy. (1) is used to treat degenerative disease, particularly cognitive disorders such as Alzheimer's disease. Sense and antisense oligonucleotides derived from the gene may be used in diagnostics and other techniques. This sequence encodes the human CHOT protein described in the invention

XX Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;

Query Match 78.9%; Score 1375; DB 5; Length 1743;

Best Local Similarity 86.8%; Pred. No. 0;

Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

```
QY 1 ATGCTTTTCATGTTGAAAGAGCTGTAGCTATTATCTCTTCTTACCTCTTAATTTCTG 60
DB 1 ATGGCTTTTCCATGTTGAAAGAGCTGTAGCTATTATCTCTTCTTACCTCTTAATTTCTG 60
QY 61 GTTGGAAATATGGGCTGCATGTGAAACCAAAAACAGGGCAACCCAGAAAGGCGAGTGA 120
DB 61 GTTGGAAATATGGGCTGCCTGAGAAACCAAAAACAGTGGCAGCGCAAGAGGCGAGCGAA 120
QY 121 GCCATCATAGTCGGGGGCGGTGACATTTGTTGTTGTTGTTGTTTAAACATGACAGCC 180
DB 121 GCCATCATAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTTAAACATGACAGCT 180
QY 181 ACCTGGGTTGAGAGGCTCACTCAATGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 ACCTGGGTTGAGAGGCTCACTCAATGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 GGTAGCTTGGGCTCATGCAACCATTTGATTTCTGAGTCTATTTTAAAGTGTG 300
DB 241 GGTAGCTTGGGCTCATGCAACCATTTGATTTCTGAGTCTATTTTAAAGTGTG 300
QY 301 TTTTTCGGAACCTATGCGTTTCAAGGATATGTACTATGTAGACCATTTCAAAACAG 360
DB 301 TTTTTCGGAACCTATGCGTTTCAAGGATATGTACTATGTAGACCATTTCAAAACAG 360
QY 361 ATCTATGAAAGCGATGGGTGGGCTGCTTTCATCCCTGCACTGATGGAGAGAGATGTC 420
DB 361 ATCTATGAAAGCGATGGGTGGGCTGCTTTCATCCCTGCACTGATGGAGAGAGATGTC 420
QY 421 TGGGCTGAGCAATTTCTCTGCTTTGGAGGCAACATGAGCTGATCATGATGTGAT 480
DB 421 TGGGCTGAGCAATTTCTCTGCTTTGGAGGCAACATGAGCTGATCATGATGTGAT 480
QY 481 GTGAACATATCGGTATGCTCTGCACTGATTTGCCATTTTAAACCTATGAGGTGG 540
DB 481 GTGAACATATCGGTATGCTCTGCACTGATTTGCCATTTTAAACCTATGAGGTGG 540
QY 541 CTCTACTCTGTGGCATATCTGATGTTGTCAGCTATTCTGCAATTTTAAATGAGCTGG 600
DB 541 CTCTACTCTGTGGCATATCTGATGTTGTCAGCTATTCTGCAATTTTAAATGAGCTGG 600
QY 601 ATCAGTGTCCCTTTGGCCCTGTCATCTGCACTGCACTGCACTGCACTGCACTG 660
DB 601 ATCAGTGTCCCTTTGGCCCTGTCATCTGCACTGCACTGCACTGCACTGCACTG 660
QY 661 CATGCTAATATCCAGAGTCCCTGGGAGACCATTTGAATCAGTTGAAGTCTACCTGG 720
DB 661 CATGCTAATATCCAGAGTCCCTGGGAGACCATTTGAATCAGTTGAAGTCTACCTGG 720
QY 721 CTGATATATTTCTGTTATGATGCTGGGTGAATCCCATGGCAAGCTTCCAGAG 780
DB 721 CTGATATATTTCTGTTATGATGCTGGGTGAATCCCATGGCAAGCTTCCAGAG 780
QY 781 GTCTCTCTGATCTCAGCCACTATGCTCAGTACTGCTTCTGCGACCTTTTGGG 840
DB 781 GTCTCTCTGATCTCAGCCACTATGCTCAGTACTGCTTCTGCGACCTTTTGGG 840
```

```
DB 781 GTTCTCTTCTTCTCTCAGCCACTATGCTCAAGTCTGTCTCTTCTGCGACCTTTGGG 840
QY 841 TGGCTGTGATGGCTCTACCCGCCATATGATAGAGCTATTTGAGCTTTCACAGACTGG 900
DB 841 TGGCTGTGATGGCTCTACCCGCCATATGATAGAGCTATTTGAGCTTTCACAGACTGG 900
QY 901 AACCGACTGCTTACGGGATATCCAGATCCCAAGACTAAGAGAGAGAGAGAGATGTC 960
DB 901 AACCGACTGCTTACGGGATATCCAGATCCCAAGACTAAGAGAGAGAGAGAGATGTC 960
QY 961 CCGATGCTTTCAGAGTACCTCTGCGCCCTGTATACCTCTCTTCTGAGCTTGTGCTGT 1020
DB 961 CCGATGCTTTCAGAGTACCTCTGCGCCCTGTATACCTCTCTTCTGAGCTTGTGCTGT 1020
QY 1021 TCACTGCTGATGATGCTCTCAGCTGATCTGATCTGATCTGATCTGATCTGATCTG 1080
DB 1021 TCACTGCTGATGATGCTCTCAGCTGATCTGATCTGATCTGATCTGATCTGATCTG 1080
QY 1081 CGGAATATCTACAGCTTCTCTTCAAGAAATGATCAGACAGAGAAATGTTGGGTC 1140
DB 1081 CGGAATATCTACAGCTTCTCTTCAAGAAATGATCAGACAGAGAAATGTTGGGTC 1140
QY 1141 ATGAGATCACTGCTGCTTGTGTTGAGAGATCTGCAACAGCCATGCTTGTGCTGAG 1200
DB 1141 ATGAGATCACTGCTGCTTGTGTTGAGAGATCTGCAACAGCCATGCTTGTGCTGAG 1200
QY 1201 ACTGTATATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
DB 1201 ACTGTATATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
QY 1261 CTGCTCTGTATCTCTTCAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 CTGCTCTGTATCTCTTCAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 TTTGACTATTTCTGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1321 TTTGACTATTTCTGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 TTTTACCTGCTTATCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1381 TTTTACCTGCTTATCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 ACTCTCTCAGAGTAACTCTGCAATCTTCAATCAATTTGTTCTTCAATGAGAGAT 1500
DB 1441 ACTCTCTCAGAGTAACTCTGCAATCTTCAATCAATTTGTTCTTCAATGAGAGAT 1500
QY 1501 CTATTTGAAAGTGAACCTTGGCTCCCAAAATTAATGATTTGATGCTGTTGTCAGAG 1560
DB 1501 CTATTTGAAAGTGAACCTTGGCTCCCAAAATTAATGATTTGATGCTGTTGTCAGAG 1560
QY 1561 CACAGTGAAGAGAGATGAGACAGATCTTCTAGTCAAGATGAATGAATTAATTAAT 1620
DB 1561 CACAGTGAAGAGAGATGAGACAGATCTTCTAGTCAAGATGAATGAATTAATTAAT 1620
QY 1621 GAACTTGAACCTGGAACCTGCGAGAGCTTAAACCTCAGTTCAACTTCAACCAATAG 1680
DB 1621 GAACTTGAACCTGGAACCTGCGAGAGCTTAAACCTCAGTTCAACTTCAACCAATAG 1680
QY 1681 GAGGCTCTCTTGAATGTTGATTCAGTCCGAGAGGGGCTGGGACTGAAGATTAATTA 1740
DB 1681 GAGGCTCTCTTGAATGTTGATTCAGTCCGAGAGGGGCTGGGACTGAAGATTAATTA 1740
QY 1741 TGA 1743
DB 1741 TGA 1743
```

RESULT 13

ADD50638 standard; cDNA; 1743 BP.

XX AC ADD50638;

DT 15-JAN-2004 (first entry)
XX cDNA encoding human high-affinity choline transporter (hCHT).
DE
XX
XX Human; high-affinity choline transporter; hCHT; chromosome 2q12;
KW cholinergic function; Parkinson's disease; Huntington's disease;
KW Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis;
KW brain; cholinergic signaling; antiparkinsonian; anticonvulsant;
KW neurotropic; neuroprotective; neuroleptic; gene; ss.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FT CDS 1.1743
FT /tag= a
FT /product= "hCHT"
XX
XX US2003114399-A1.
XX
XX 19-JUN-2003.
XX
XX 23-JUL-2001; 2001US-00911077.
XX
XX 23-JUL-2001; 2001US-00911077.
XX
XX (BLAK/) BLAKELY R D.
XX (APP/) APPARISUNDARAM S.
XX (FERG/) FERGUSON S.
XX
XX Blakely RD, Apparsundaram S, Ferguson S;
XX
XX MPI; 2003-810914/76.
XX P-PSDB; ADD50639.
XX
XX Novel isolated polynucleotide encoding human or mouse high affinity
XX choline transporter polypeptide, useful in gene therapy to increase
XX cholinergic function in a cell of a patient suffering from Alzheimer's
XX disease.
XX
XX Claim 2; SEQ ID NO 1; 74pp; English.
XX
XX The present invention relates to the isolation of polynucleotide
XX sequences encoding human and mouse high-affinity choline transporter
XX (hCHT and MCHT respectively), and the proteins they encode. The gene
XX encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
XX encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT
XX polynucleotide sequence when delivered to a cell, increases cholinergic
XX function in the cell that is in a patient having Parkinson's disease,
XX Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
XX myasthenia gravis. The hCHT antibody is useful for controlling
XX transporter CHT proteins to the brain, and for treating the above
XX mentioned diseases. The antibody is also useful for diagnosing the above
XX mentioned disorders and to detect the influence of cholinergic
XX signaling. The present sequence encodes hCHT. Note: The sequence data
XX for this patent was obtained in electronic format directly from the USPTO
XX web site at seqdata.uspto.gov.
XX
XX Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;
XX
XX Query Match 78.9%; Score 1375; DB 10; Length 1743;
XX Best Local Similarity 86.8%; Pred. No. 0;
XX Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

DB 121 GCCATCATAGTGTGGGCGAGATATGGTTATTGGTTGGTGGATTTACATGACGCT 180
QY 181 ACCTGGGTGGAGAGAGGCTACATCATATGGAGACGAGAGAGCTATGGCCAGGTTGT 240
DB 181 ACCTGGGTGGAGAGAGGCTATATCATATGGAGAGGCTATATGGAGAGGTTAT 240
QY 241 GGCTATGCTTGGGCTCATGACCATCATATGATTTCTGAGTCTAAATTTAGTGGTCTG 300
DB 241 GGCTATGCTTGGGCTCATGACCATCATATGATTTCTGAGTCTAAATTTAGTGGTCTG 300
QY 301 TTTTTCGAAACCTATATGCGTTCCAAAGGATATGACATATGTTAACCATTCAAAACAG 360
DB 301 TTTTTCGAAACCTATATGCGTTCCAAAGGATATGACATATGTTAACCATTCAAAACAG 360
QY 361 ATCTATGAAAGGAGATGGGCTGGCTCTTCATGCTGACCTGATGGAGAGATGTTTC 420
DB 361 ATCTATGAAAGGAGATGGGCTGGCTCTTCATGCTGACCTGATGGAGAGATGTTTC 420
QY 421 TGGGCTGACGAATTTCTCTGCAATTAAGGAGCCACCATGAGCGTGATCATATGATGAT 480
DB 421 TGGGCTGACGAATTTCTCTGCAATTAAGGAGCCACCATGAGCGTGATCATATGATGAT 480
QY 481 GTGAACATATCGGTCATGTTCTCTGCACTGATTCGCAATTTCTTAATACCTAGTGGTGG 540
DB 481 GTGAACATATCGGTCATGTTCTCTGCACTGATTCGCAATTTCTTAATACCTAGTGGTGG 540
QY 541 CTCTACTCTGTCGCAATATGATGTTGTCAGCTATTCGCAATTTTAAAGAGCTGTG 600
DB 541 CTCTACTCTGTCGCAATATGATGTTGTCAGCTATTCGCAATTTTAAAGAGCTGTG 600
QY 601 ATCAGTGTCCCTTTTGGCCCTGTCACATCCCTGACATCCGACACGACATCGGATTCACAGCTGTG 660
DB 601 ATCAGTGTCCCTTTTGGCCCTGTCACATCCCTGACATCCGACACGACATCGGATTCACAGCTGTG 660
QY 661 CATGCTAAATACGAGAGTCCCTGCTGGTGGAGAACATTTGAATCACTGATGCTTACCTGTG 720
DB 661 CATGCTAAATACGAGAGTCCCTGCTGGTGGAGAACATTTGAATCACTGATGCTTACCTGTG 720
QY 721 CTGTATATTTTCTGTATATGATGCTGGGATGGAATCCCAAGGACGCTACTCCAGAGG 780
DB 721 CTGTATATTTTCTGTATATGATGCTGGGATGGAATCCCAAGGACGCTACTCCAGAGG 780
QY 781 GTCTCTCTTCAATCCCAAGGACGCTATGCTGAGTGTCTCTTCTGAGCTTTTGGG 840
DB 781 GTCTCTCTTCAATCCCAAGGACGCTATGCTGAGTGTCTCTTCTGAGCTTTTGGG 840
QY 841 TGGCTGTATGAGCTCTACCCGCAATATGATATGAGCTATTTGAGCTTCCACAGACTGG 900
DB 841 TGGCTGTATGAGCTCTACCCGCAATATGATATGAGCTATTTGAGCTTCCACAGACTGG 900
QY 901 AACCAAGCTGCTGAGGATGCAATCCCAAGCTATGAGGAGAGACAGATGATTTCTC 960
DB 901 AACCAAGCTGCTGAGGATGCAATCCCAAGCTATGAGGAGAGACAGATGATTTCTC 960
QY 961 CCGATGCTTCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 CCGATGCTTCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 TCAAGCTGCTGATATGCTGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 TCAAGCTGCTGATATGCTGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 CGGAATATATACAGCTTCTTCAAGAAATATGATCAAGAAAGAAATATGTTGGGCT 1140
DB 1081 CGGAATATATACAGCTTCTTCAAGAAATATGATCAAGAAAGAAATATGTTGGGCT 1140
QY 1141 ATGAGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1141 ATGAGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 ACTGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 ACTGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260

Db	601	ATCAGCGTCCCTTTGCAATTGTCAATCTGCAATCGGACGATCGGGTTACCTGCTGTG	660
Qy	661	CATGCTAAATACCAAGTCCCTGGCTGGGAACAATTGAATCATTTAAATCAACCTGG	720
Db	661	CATGCAAAATACAAAAGCCGTGGCTGGGAATGTGATCATCATGAAAGTCACTCTTGG	720
Qy	721	CTGTAAATATTTCTGTTATTTGATGTGGGAGAAATCCATGGCAAGCTACTTCACAGG	780
Db	721	CTGTATATTTTCTGTGTTTGTATGTCTGGGTGAAATCCATGGCAAGCATCTTTCAAGG	780
Qy	781	GTCCTCTCTTCATCCTCAAGCCATATGCTCAGGATCTGTCTTCCTGGCAGCTTTTGG	840
Db	781	GTTCTCTCTTCTCTCCTAGCCACTATGCTCAAGTGCTGTCTTCCTCGGACGCTTTGGG	840
Qy	841	TGCCTGGGATAGGCTTACCCGCCATATGCAATAGAGTATTTGGAGCTTCCACAGACTGG	900
Db	841	TGCCTGGGATAGGCTATCCAGCCATATCCAGCATCTATGGGGCATTTGAGAGCATCAACAGACTGG	900
Qy	901	AACCAAGACTGCTCAAGGGTATCCAGATCCCAAGACTAAAGAGAGAGACATGATCTC	960
Db	901	AACCAAGACTGATATGGGCTTCAGATCCCAAGACTACAGAAAGGACAGATGATTTTA	960
Qy	961	CGATTCGTTCTGCAGTACTCTGCCCTGTGTACATCTCCTCTTTGGGCTTGGTCTGT	1020
Db	961	CCAATTTGTTCTGCAGTATCTCTGCCCTGTGTATTTCTTCTTTGGCTTGGGACGT	1020
Qy	1021	TCAGCTGTGTCAATGCTCCTCAGCTGACCTGCTCATCTGCTGGGGAGTTCTATGTTGCT	1080
Db	1021	TCGTGTCTGTATATGTCATCAGCAATTTCTTCACATCTGTAGCAAGTTCCATGTTTGA	1080
Qy	1081	CGGAATATCTACCAAGCTTTCCTTCAGACAAATATGCATCAGCAAGAAATGTGTGGGCT	1140
Db	1081	CGGAATATCTACCAAGCTTTCCTTCAGACAAATATGCTTCGACAAAGAAATGTTTGGGTT	1140
Qy	1141	ATGAGGATCACTGTGCTGTGTGTGGAGCATCTGCAACGCCATGCTTGTCTACGAG	1200
Db	1141	ATGGAATCAACAGGTGTGTGTGTGGAGCATCTGCAACGCCATGCGCTGTCTACGAAA	1200
Qy	1201	ACTGTGATAGGGCTCTGGTACTAGAGCTCTAGCCTGTGCTCATCATCATCTTCCACAG	1260
Db	1201	ACTGTGATAGGGCTCTGTGACTCATGTTCTGACCTTGTTCATGTATATCTTCCCCAG	1260
Qy	1261	CTGCTGTGTGACTCTTCATCAAGGAACCAACCTTATGGGGGAGTTGCTGTATATTT	1320
Db	1261	CTGCTGTGTGACTCTTGTATTAAGGAACCAACCTATGGGGGCTGTGGAGTTATGTT	1320
Qy	1321	TTTGGACATATCCGAGAAATTACGTGGAGAGACCATATCTATCTTGACGCCCTTATC	1380
Db	1321	TCGTGGCTCTTCTCGAGAAATACGTGAGGGAGCCATATCTGTATCTTCAGCCCTTATC	1380
Qy	1381	TTCTTACCTGGTTTATCTCTGACAGAGATGTATATACATTCAGAGGTTCCATTTAAA	1440
Db	1381	TTCTTACCTGGCTTTTACCTCGATGATATGTGTATATATATCAAGAAATTTCCATTTAAA	1440
Qy	1441	ACTCTTCATAGTTACTCTCATTTCTTTACCAACATTTGTGTCTTATCTTACCCAAGTAT	1500
Db	1441	ACACTTGGCAATGTTACATCAATCTTAAACCAACATTTGCAATCTCTATTCAGCCAAAGTAT	1500
Qy	1501	CTTATTTGAAATGGAACCTTGCTCCCAAAAATTAAGTATTTGATGCTGTGCGCAGAG	1560
Db	1501	CTTATTTGAAATGGAACCTTGCTCCACCTTAAATTAATGATTTGATGCTGTGTGCGAAGA	1560
Qy	1561	CACAGTGAAGAAATATGGAACAAGCCATTTAGTCAGAAATGAAAATATCAAAATTAAT	1620
Db	1561	CACAGTGAAGAAAATATGGAATAGCAATCTTGTCAAAAATGAAAATATTAATTTGAT	1620
Qy	1621	GAACTTGAACCTGAGAAACCTCGGAGAGGCTTAACCTCAGTTCAACCTTACCAAAATAG	1680
Db	1621	GAACTTGAACCTTGTGAAGCCACGACAGAGCATGACCTCAGCTCAACCTTCAACCAATAA	1680
Qy	1681	GAGGCGCTCTTGATGTGATTTCAAGTCGGAGGGGCTGTGGAGCTGAAGATTAATTTACAA	1740

[illegible]

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: July 10, 2006, 14:02:30 ; Search time 855 Seconds
(without alignments)
7094.569 Million cell updates/sec

Title: US-10-724-806-4

Sequence: 1 MFPHVGHVAILFYLIIFL.....FALLDVDSPEGGSTEDNLQ 580

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues
Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DEV=xlp
-O=/abs/ABSWEB/spool/US10724806/runat.10072006.140239.7586/app-query.fasta.1
-DB=N.GeneSeq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs505p
-USER=US10724806 @CCN 1 1.1147 @runat.10072006.140239.7586 -NCRU=6 -ICPU=3
-NO MMAP -NRC SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N GeneSeq 8:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2004as:*
- 12: geneseqn2004bs:*
- 13: geneseqn2004cs:*
- 14: geneseqn2005s:*
- 15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2993	100.0	1743	10	ADD50640 CDNA enco
2	2993	100.0	1743	10	ADD50660 CDNA enco
3	2993	100.0	1743	14	ADV77898 Mouse hig

4	2993	100.0	1743	14	ADV77918
5	2968	99.2	1743	4	AAFB1713
6	2967	99.1	4938	5	AAAD02457
7	2956	98.8	1743	4	AAFB1711
8	2956	98.8	4904	10	ADD50642
9	2956	98.8	4904	14	ADV77900
10	2795	93.4	1743	4	AAFB1712
11	2795	93.4	1743	5	AAH49207
12	2795	93.4	1743	9	ABX94338
13	2795	93.4	1743	10	ADD50638
14	2795	93.4	1743	14	ADV77896
15	2795	93.4	1813	10	ADD50646
16	2795	93.4	1813	14	ADV77904
17	1491.5	49.8	1833	12	ADU93831
18	1444.5	48.3	1731	4	AAFB1710
19	1444.5	48.3	1985	10	ADD50644
20	1444.5	48.3	1985	14	ADV77902
21	1413	47.2	4223	4	ABL29568
22	1405	46.9	1729	6	ABL29569
23	1099.5	36.7	1461	6	ABK73210
24	1096	36.6	8760	5	AAH49206
25	1096	36.6	26323	14	AEB16466
26	1096	36.6	119040	10	ADD50656
27	1096	36.6	119040	14	ADV77914
28	1096	36.6	142299	10	ADD50651
29	1096	36.6	142299	14	ADV77909
30	601	20.1	386	5	ADP02461
31	544	18.2	329	12	ADP08298
32	342	11.4	349980	5	AAFB6431
33	318	10.6	2475	6	AB212050
34	318	10.6	2475	12	ADMA4568
35	313	10.5	2028	6	ABL57735
36	313	10.5	2105	8	ABT33401
37	313	10.5	2105	12	ADH42364
38	313	10.5	2105	12	ADH42358
39	313	10.5	2200	4	AAE27713
40	313	10.5	2384	13	ADOC89141
41	313	10.5	2384	13	ADH44927
42	312	10.4	2028	4	AAFB4022
43	312	10.4	2456	4	AAFB4023
44	310	10.4	2326	6	ABK10764
45	307	10.3	2043	13	ADOC89049

ALIGNMENTS

RESULT 1
ADD50640
ID ADD50640 standard; CDNA; 1743 BP.
XX
AC ADD50640;
XX
DT 15-JAN-2004 (first entry)
XX
DE CDNA encoding mouse high-affinity choline transporter (mCHT) #1.
XX
KW Mouse; high-affinity choline transporter; mCHT; cholinergic function;
XX
KW Parkinson's disease; Huntington's disease; Alzheimer's disease;
XX
KW schizophrenia; dysautonomia; myasthenia gravis; brain;
XX
KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;
XX
KW neuroprotective; neuroleptic; gene; ss.
XX
OS Mus sp.
XX
FH
FT Key Location/Qualifiers
FT CDS 1..1743 /*tag= a
FT /product= "mCHT #1"
XX
PN US2003114399-A1.
XX
PD 19-JUN-2003.
XX

PF 23-JUL-2001; 2001US-00911077.
XX 23-JUL-2001; 2001US-00911077.
XX 23-JUL-2001; 2001US-00911077.
XX (BLAK/) BLAKELY R D.
PA (APPA/) APPASUNDARAM S.
PA (FERG/) FERGUSON S.
XX
XX Blakely RD, Appasundaram S, Ferguson S;
XX MPI: 2003-810914/76.
DR P-PSDB; ADD50641.
XX
XX Novel isolated polymucloctide encoding human or mouse high affinity
PT choline transporter polypeptide, useful in gene therapy to increase
PT cholinergic function in a cell of a patient suffering from Alzheimer's
PT disease.
XX
XX Claim 30; SEQ ID NO 3; 74pp; English.
XX
XX The present invention relates to the isolation of polymucloctide
CC sequences encoding human and mouse high-affinity choline transporter
CC (hCHT and mCHT respectively), and the proteins they encode. The gene
CC encoding hCHT is located on chromosome 2q12. The polymucloctide sequence
CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT
CC polymucloctide sequence when delivered to a cell, increases cholinergic
CC function in the cell that is in a patient having Parkinson's disease,
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
CC myasthenia gravis. The hCHT antibody is useful for controlling
CC transporter:CHT proteins to the brain, and for treating the above
CC mentioned diseases. The antibody is also useful for diagnosing the above
CC mentioned disorders and to detect the influence of cholinergic
CC signaling. The present sequence encodes mCHT. Note: The sequence data
CC for this patent was obtained in electronic format directly from the USPTO
CC web site at seqdata.uspto.gov.
XX
XX Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 6.05e-314 Length: 1743
Score: 2993.00 Matches: 580
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0
US-10-724-806-4 (1-580) x ADD50640 (1-1743)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrIleuLeuIlePheLeu 20
DB 1 ARGCCCTTCCATGTGGAGAGACTGTAGCTATTATCTCTTCACTCTTATATTCTG 60
QY 21 ValGlyIleTrrAlaAlaTrrPlysThrIlyAsnSerGlyAsnProGluGluArgSerGlu 40
DB 61 GTTGGAAATGGGCTGCATGAGAAACCAAAACAGGGGCAACCCAGAAAGGCCATGAA 120
QY 41 AlaIleIleValAlGlyIlyArgAspIleGlyLeuLeuValGlyIlyPheThrMetThrAla 60
DB 121 GCCATCATATGTGGGGGGCGGTGACATTTGTTGTTGGTGGTTTACCATGACAGCC 180
QY 61 ThrTrrPValGlyIlyGlyTrrIleAsnGlyThrAlaGluAlaValTrrGlyProGlyCys 80
DB 181 ACCTGGGTGGAGAGAGCTACATCAATGGAGACAGAGCAAGCTGTATGGGCGAGTTGT 240
QY 81 GlyLeuAlaTrrAlaHisAlaProlIleGlyTrrSerLeuSerLeuIleLeuGlyIlyLeu 100
DB 241 GGCTTACGCTTGGGCTATGACCCCATTTGATTTCTGAGTCTATATTTTAAAGTGAGTCTG 300
QY 101 PhePheAlaIlyProMetArgSerIlySGlyTrrValThrMetLeuAspProPheIlySgl 120
DB 301 TTTTTCGCAAAACCTATGGCTTCCAGAGGATATGTGACATATGTTAACCATTCAGAACG 360
QY 121 IleTrrGlyIlyLeuArgMetGlyIlyLeuLeuPheIleProAlaLeuMetGlyIlyMetPhe 140

DB 361 ATCTATGGAAAGCCGATGGGTGGCTGCTCTTCAATCCCTGCATGATGGAGAGATGTTTC 420
QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyValaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGCAGCAATTTTCTCTGCATTTAGGGGCCACACATCAGCGATCATTTGATGTGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTrrThrLeuValGlyIly 180
DB 481 GTGAACATATGCGCATGTCTCTGCACCTATGGCATTTGTCATTATACCTTATGCTGGTGG 540
QY 181 LeuTrrSerValAlaTrrThrAspValaGluLeuPheCysIlePheIleGlyLeuTrrP 200
DB 541 CTCTACTCTGGGCAATATGATGATGTTTCCAGCTATTCGATTTTATAGACTGTGG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB 601 ATCATGTGCCCTTTTGGCTGTCATCTCGACATCTCGACATCGCATCGGATTCACAGCTGTG 660
QY 221 HisAlaIlyTrrGlnSerProTrrPleuGlyThrIleGlySerValGluValTrrThrTrp 240
DB 661 CATGCTAATATCCAGAGTCCCTGGGGAACCATTTGAATCAGTTGAAGTCTACACCTGG 720
QY 241 LeuAspAsnPheLeuLeuMetLeuGlyIlyIleProTrrPglAlaTrrPheGlnArg 260
DB 721 CTGTAAATTTTCTGTATATGATGCTGGGTGGAATCCATGCAAGCTTACTTCAGAGG 780
QY 261 ValLeuSerSerSerSerAlaThrTrrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 781 GTCCCTCTTATCTTCACTTCAGCAACCTATGCTCAGGACTGCTCTTGGAGCTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrSptrr 300
DB 841 TGCCCTGGATGGCTCTACCCGCCATATGATAGAGAGCTATTTGAGCTTCCACAGCTGG 900
QY 301 AsnGlnThrAlaTrrGlyIlyTrrProAspProlIlySglGluAlaAspMetIleLeu 320
DB 901 AACCAAGACTGCTTACGGGATTCAGATCCAGACTTAAGAGAAAGCAAGATGATTTCTC 960
QY 321 ProlIleValLeuGlnTrrLeuGlyProValTrrIleSerPhePheGlyLeuGlyAlaVal 340
DB 961 CCGATCGTTTGCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1021 TCAGCTGCTGTCATGCTCTCAGCTGACCTGATCTGTCATCTGTCGAGGATTTGTTGCT 1080
QY 361 ArgAsnIleTrrGlnLeuSerPheArgGlnAsnAlaSerAspIlySglIleValTrrVal 380
DB 1081 CGGAATATCTACAGCTTTCTTTCAGACAAATGCAATGCAACAGAAATTTGTGGGTTC 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIly 400
DB 1141 ATGAGAGATCAGTGTCTGTTGGAGAGATCTGCAGAGCATGCGCTTGTGTCAGAGAG 1200
QY 401 ThrValIlyTrrGlyLeuTrrTrrLeuSerSerAspLeuValTrrIleIleIlePheProGln 420
DB 1201 ACTGTGTATGGGCTCTGAGACTGAGCTCTGACCTTGTCTCATCATATCATCTTCCACAG 1260
QY 421 LeuLeuCysValLeuPheIleIlySglIlyThrAsnTrrTrrGlyAlaValAlaIlyTrrIle 440
DB 1261 CTGCTCTGTATCTCTTCAACAGAAACCAACTTATGGGGAGCTTGTGTTATATTT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyIlyGluProTrrLeuTrrLeuGlnProLeuIle 460
DB 1321 TTTGGAATTTCTCTGAGATTTCTGAGAGAGGCCATATCTATCTTGAGCCCTTAATC 1380
QY 461 PheTrrProGlyTrrTrrSerAspIlySasngIlyIleTrrAsnGlnArgPheProPheIly 480
DB 1381 TTCTACCTGGTATTAATCTGACAAAGATGTATATCAACAGAGGTTCCCATTTTAA 1440
QY 481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTrrLeuAlaIlyTrr 500

Db	1441	ACTCTCTCCAAAGTTACCTCATTTCTTTCACCAAACTTGTGTCTTACTAGCCCAAGTAT	1500
Qy	501	LeuPheGluSerGlyIleuLeuProProIleuAspValPheAspAlaValAlaIleArg	520
Db	1501	CTATTGTGAAGATGGAACTCTCCCTCCCAAAATTAAGTGAATTTGAAGCGCTGTGTGCAAG	1560
Qy	521	HisSerGluGluAsnMetAspIleThrIleLeuValArgAsnGluAsnIleIysLeuAsn	540
Db	1561	CACAGTGAAGGAACATGAGACAAAGCCATTCTAGTCGAAATGTAAATTCAAATTAAAT	1620
Qy	541	GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys	560
Db	1621	GAATTTGACCACTGTGAAACCTCGCGAGAGCCTAACCTCAATTCAACATTATAG	1680
Qy	561	GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyIleuGluAspAsnLeuGln	580
Db	1681	GAGGCGCCCTCTGTGAATGTGATTTCCAGTCCGAGGGGGTCTGGAGCTGAAGAATTAATTTACA	1740

XX	RESULT 2
XX	ADD50660
ID	ADD50660 standard; cDNA, 1743 BP.
AC	
XX	ADD50660;
DT	15-JAN-2004 (first entry)
XX	
DE	cDNA encoding mouse high-affinity choline transporter (mCHT) #2.
XX	
KW	Mouse; high-affinity choline transporter; mCHT; cholinergic function; Parkinson's disease; Huntington's disease; Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis; brain; cholinergic signaling; antiparkinsonian; anticonvulsant; nootropic; neuroprotective; neuroleptic; gene; ss.
XX	
OS	Mus sp.
XX	
FH	Location/Qualifiers
FT	1..1743
FT	/tag= a
FT	/product= "mCHT #2"
XX	
PN	US2003114399-A1.
XX	
PD	19-JUN-2003.
XX	
PF	23-JUL-2001; 2001US-00911077.
XX	
PR	23-JUL-2001; 2001US-00911077.
XX	
PA	(BLAK/) BLAKELY R D. (AAPA/) APPARISUNDARAM S. (FERG/) FERGUSON S.
XX	
PI	Blakely RD, Apparsundaram S, Ferguson S;
XX	
DR	WPI: 2003-810914/76.
DR	P-PSDB; ADD50661.
XX	
PT	Novel isolated polynucleotide encoding human or mouse high affinity choline transporter polypeptide, useful in gene therapy to increase cholinergic function in a cell of a patient suffering from Alzheimer's disease.
PT	
XX	
PS	Example 4; SEQ ID NO 23; 74bp; English.
XX	
CC	The present invention relates to the isolation of polynucleotide sequences encoding human and mouse high-affinity choline transporter (hCHT and mCHT respectively), and the proteins they encode. The gene encoding hCHT is located on chromosome 2q12. The polynucleotide sequence encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT polynucleotide sequence when delivered to a cell, increases cholinergic function in the cell that is in a patient having Parkinson's disease, Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or

myasthenia gravis. The hCMT antibody is useful for controlling transporter CMT proteins to the brain, and for treating the above mentioned disorders. The antibody is also useful for diagnosing the above mentioned disorders and to detect the influence of cholinergic signaling. The present sequence encodes mCMT. Note: The sequence data for this patent was obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov.

US-10-724-806-4 (1-580) X ADD50660 (1-1743)

QY	1	MetProPheHISValGIuGIyLeuValAlaIleIleLeuPheThyLeuLeuIlePheLeu	20
Db	1	ATGCTTTCACATGGGAAGAGCATGGTAGCTATTATCTCTTCTACCTCTCTAATTTCTG	60
QY	21	ValGIyIleTPAlaAlaIleTPHyThrIysAsnSerGIyAsnProGIuGIuArgSerGIu	40
Db	61	GTTGGAAATATGGGCTGCATCGAAAAACAAAAACAGGGGCAACCAAGAGCGCAGTGA	120
QY	41	AlaIleIleValGIyGIyArgAspIleGIyLeuLeuValGIyGIyPheThrMetThra	60
Db	121	GCCATCATAGTCGGGGGCGGTGACATTTGGTTGGTTGGTTGGTTTTCATGACAGACC	180
QY	61	ThrTPValGIyGIyGIyTrpIleAsnGIyThraIleValIleValGIyGIyProGIyCys	80
Db	181	ACCTGGGTGGAGAGGGCTACATCATAGGGAACAAGAACAGTATAGGGCCAGGTGT	240
QY	81	GIyLeuAlaTPAlaHISAlaProIleGIyTrpSerLeuSerLeuIleLeuGIyGIyLeu	100
Db	241	GGCTAGCTTGGGGCTCATGCACCATTTGGATATTCTCTAGTCTAATTTTAGGTGGTCTG	300
QY	101	PhePheAlaIysProMetArgSerIysGIyTrpValThMetLeuAspProPheIysGln	120
Db	301	TTTTTGGCAACCTATGCGTTCCAAAGGATATGTGACTATGTTAAGCCATTCAACAG	360
QY	121	IleTrpGIyIysArgMetGIyGIyLeuLeuPheIleProAlaLeuMetGIyGIyMetPhe	140
Db	361	ATCTATGGAAGCGCATGGGTGGGCTGCTTTCATCTCTGCACATGATGGAGAGATGTTT	420
QY	141	TrpAlaAlaAlaIlePheSerAlaLeuGIyAlaThrIleSerValIleIleAspValAsp	160
Db	421	TGGCTGCAGCAATTTCTCTGCATTATGGGGCCACCATCAGCGTGATCATGTATGTGAT	480
QY	161	ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTrpThrLeuValGIyGIy	180
Db	481	GTTGAACATATGGGTCAATGTCTCTGCACATTCATCTTATATCCCTAGGGGGTGGG	540
QY	181	LeuTrpSerValAlaIleTrpThrAspValValGIyLeuPheCysIlePheIleGIyLeuTrp	200
Db	541	CTTAACTGTGGCAATATCATGATGTGTCCAGCTATTTGTGCATTTTATAGGACGTGG	600
QY	201	IleSerValProPheAlaLeuSerHISProAlaValThraSpIleGIyPheThrAlaVal	220
Db	601	ATCACTGTCTCTTTGGCCCTGTCAATCTCGCAGTACCGGACATCGGATTCACAGCTGTG	660
QY	221	HISAlaIleTrpGlnSerProTrpLeuGIyThrIleGlnSerValGIuValTrpThrTrp	240
Db	661	CATGCTAAATATCCAAAGTCCCTCGGGTGGGAACATTGAATCAGTTAAGCTAACACTGG	720
QY	241	LeuAspAsnPheLeuLeuMetLeuGIyGIyIleProTrpGlnAlaTrpPheGlnArg	260
Db	721	CTTGTAATATTTCTGTATATGATGCTGGGTGGAAATCCATGGCAAGCTTATCCAGAGG	780
QY	261	ValLeuSerSerSerAlaThrTrpAlaGlnValLeuSerPheLeuAlaIlePheGIy	280

Db	781	GTCTCTCTTATCTCTCAGCCACCTATGCTCAGTACGTCCTTCCTGCGACTTTGGG	840
Qy	281	CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp	300
Db	841	TGCCTGGTATGAGCTCTACCCGCCATATGACATAGAGGATATTGGAGCTTCCACAGACTGG	900
Qy	301	AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluAlaAspMetIleLeu	320
Db	901	AAACAGACTGCTTACGGGATTCACATCCCAAGACTTAAGAGGAAGCAGATATTTCTC	960
Qy	321	ProIleValLeuGlnTyrLeuCysProValTyrIleSerPheGlyLeuGlyAlaVal	340
Db	961	CCGATGCTTCTCAGACTCTGCCCTGTGTACATCTCTTCTTGGGCTGGTGGT	1020
Qy	341	SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla	360
Db	1021	TGAGCTGCTGATATGCTCCACACTGATCTGTCAATCTGTGGCGAGTTCTATGTTGCT	1080
Qy	361	ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrpVal	380
Db	1081	CGGAATATTCACACGCTTTCCTTCAGACAAATGATCAGACAAAGAAATTGTGGGCTG	1140
Qy	381	MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys	400
Db	1141	ATGAGGATCATCGTGTCTGTGTTCGAGAGATCTGCAGACCATGCTTCTGCGAG	1200
Qy	401	ThrValTyrGlyLeuTrpTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln	420
Db	1201	ACTGTATAGGGGCTCTGGTACTGAGCTTGACCTTGTCTACATATATCTTCCACAG	1260
Qy	421	LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle	440
Db	1261	CTGCTCTGTGTACTCTTCATCAAGAACCAACACTTATGGGGCGTGTGGTATATT	1320
Qy	441	PheGlyLeuPheLeuArgIleThrGlyGlyLysProTyrLeuTyrLeuGlnProLeuIle	460
Db	1321	TTTGGATATCTCTCGAATTACTGGAGGAGGCCATCTATCTTCCAGCCCTTAATC	1380
Qy	461	PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys	480
Db	1381	TTCTACCTCGGTATTACTCTGACAAAGATGGTATATACATCAAGGTTCCATTAA	1440
Qy	481	ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr	500
Db	1441	ACTCTCTCCATGGTAACTTCATCTTTTACCAACATTGTGTTCTTATCTAGCCAAAGTAT	1500
Qy	501	LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValValAlaArg	520
Db	1501	CTATTGAAAGTGGAACCTTGCTCCCAAATTAGATGATTTGATGCTGTGTGCGCAGG	1560
Qy	521	HisSerGluGlnAsnMetAspLysThrIleLeuValArgAsnGlnAsnIleLysLeuAsn	540
Db	1561	CACAGTAGAGAGAACTATGACAAGACCAATTCATAGTCAGAAATGAAAAATTAAT	1620
Qy	541	GlnLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys	560
Db	1621	GAACTTGACCTGTGAACCTTGCGCAGACCTTACCTTAGTTCAACTTCAACAAATAG	1680
Qy	561	GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGlnAspAsnLeuGln	580
Db	1681	GAGGCGCTCTTGATGATTGATTCAGATCCGGAAGGGGCTGGGACGTGAAGATATTTACA	1740
RESULT 3			
ID ADV77898 standard; cDNA, 1743 BP.			
XX ADV77898;			
XX 24-FEB-2005 (first entry)			
XX Mouse high affinity choline transporter (mCHT) cDNA.			
XX			

KM	Choline transporter; neuromuscular disorder;
KM	Autonomic nervous system disorder; central nervous system disorder;
KM	Parkinsons disease; Huntingtons chorea; genetic disorder;
KM	Alzheimers disease; degeneration; neurological disease; schizophrenia;
KM	psychiatric disease; myasthenia gravis; immune disorder; gene therapy;
KM	cns-gen.; nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KM	musclar-gen.; neuroleptic; DNA purification; gene; ss.
XX	
OS	Mus musculus.
XX	
Key	Location/Qualifiers
FT	1..1743
FT	CDS
FT	/*tag= a
FT	/product= "Mouse CHT protein"
XX	
PN	US2004248838-AI.
XX	
PD	09-DEC-2004.
XX	
PF	01-DEC-2003; 2003US-00724806.
XX	
PR	23-JUL-2001; 2001US-00911077.
XX	
PA	(UYVA-) UNIV VANDERBILT.
XX	
P1	Blakely RD, Apparsundaram S, Ferguson S;
XX	
DR	WPI: 2005-020586/02.
DR	P-PSDH; ADV77899.
DR	GENBANK; AF276872.
XX	
PT	New human high affinity choline transporter (CHT) cDNA, useful for treating a neuromuscular, autonomic or central nervous system disorder, including Parkinson's disease, Huntington's disease, Alzheimers, schizophrenia.
PT	
XX	
PS	Claim 30; SEQ ID NO 3; 73pp; English.
XX	
CC	The present invention provides polynucleotides encoding novel high affinity choline transporters (CHTs), methods for their use in screening disorders, autonomic or central nervous system disorders such as Parkinsons disease, Huntingtons disease, Alzheimers disease, schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is also useful in gene therapy. The present sequence is mouse high affinity choline transporter (mCHT) cDNA. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at seqdata.uspto.gov/sequence.html?DocID=US20040248838.
CC	
CC	
XX	
SQ	Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	6.05e-314 Length: 1743
Score:	2993.00 Matches: 580
Percent Similarity:	100.0% Conservative: 0
Best Local Similarity:	100.0% Mismatches: 0
Query Match:	100.0% Indels: 0
DB:	14 Gaps: 0
US-10-724-806-4 (1-580) x ADV77898 (1-1743)	
OY	1 MetProPheHisValGlnGlyLeuValAlaIleIleuPheTyrlLeuIleuIlePheIeu 20
Dd	1 ATGCCCTTCACAGTGGAAGACTGGAGCTATATCCTCTCACTCCTTAATTTCGTG 60
OY	21 ValGIyIEtRPAlaAlaIatRPyThrLysAsnSergIyaSnProJugIuaYgSerGlu 40
Dd	61 GTTGGAAATATGGCGCTCATGAAAAACAATAACGCGCAACCAGAAGAGCGCATGAA 120
OY	41 AlaIleIleValaIGLyIaRGaspIleGlyLeuLeuValaIGLyIaPhetlMerThAla 60
Dd	121 GCACATATAGTCCGGGGCGGTATACATTGGTTGGTTGGTGGTATTAACATGACAGCC 180

QY ThrProValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
 DB ACCTGGGTTGGAGAGAGCTACATCAATGGAGACAGACAGATGTAAGGCCAGGTTGT 240
 QY GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
 DB GGCTAGCTGGGCTCATGACCACTGGATATTCTGTAGCTTAATTTAGGTGGCTCG 300
 QY PhePheAlaLysProMetLysSerGlyTyrValIleMetLeuAspPropheLysGln 120
 DB TTTTTCGCAAAACCTATGCGCTTCCAAAGGATATGATGATATGTAACCATTTCAACAG 360
 QY IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140
 DB ATCTATGAAAGGCGATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGAGAGATGTTT 420
 QY TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
 DB TGGGCTGCAGCAATTTTCTCTGCAATTAGGGCCACCATCGATGATCATTTGATGGAT 480
 QY ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
 DB GGAAACATATCGGTCAATGTCTCTGCACTCATTCATTCCTTAATCCCTAGTGGGTGG 540
 QY LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200
 DB CTCTACTCTGTGCAATATCAATGATGTTGTCACACTTCTGCAATTTTATAGACGTGG 600
 QY IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
 DB ATCAGAGTCCCTTTTCCCTGTCACATCCGACAGTCAACGACATCGGATTCACAGCTGG 660
 QY HisAlaLysTyrGlnSerProTyrPleuGlyThrIleGlySerValGluValTyrThrTyr 240
 DB CATGCTCAATATACAGAGTCCCTGGCTGGGAACCATTTGAATCAAGTTCACACTGG 720
 QY LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArg 260
 DB CTTGATTAATTTTCTGTTATGATGCTGGGTGGAAATCCCATGGACGCTTATCCAGAGG 780
 QY ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
 DB GTCCTCTCTTCAATCCACGACCATTAATGCTCAGATCTGCTTCCGAGGCTTTTGGG 840
 QY CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300
 DB TGCCTGTGATGCTCTACCCCGCATATGATGAGCTATTGGAGCTTCCACAGACTGG 900
 QY AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGlnGluAlaAspMetIleLeu 320
 DB AACCAACACTGCGCTACGGGATCCAGATCCCAAGACTTAAGAGAGAGACATGATTTCT 960
 QY ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
 DB CGGATGCTTCTGAGTACCTCTGCTGCTGCTGCTGCTGCTTCTTGGGCTTGGTGTCTT 1020
 QY SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
 DB TCAGCTGCTGTCTATGCTCCAGCTGACTGCTCATCTGTGCGGAGTTCTATGTTTGGCT 1080
 QY ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrPval 380
 DB CGGAATATCTACAGCTTCTCTTCCAGCAAAATGCAATCAGACAAAGAAATGTGTGGTTC 1140
 QY MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
 DB ATAGAGATCACTGCTGTGTGTGCGAGCATCTGCAACACCATGCTTGTGTCAGCAAG 1200
 QY ThrValTyrGlyLeuTyrPyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
 DB ACTGTATATGGGCTGTGTGATCCTGAGCTGTGACCTTGTCTACATCATCATCTTCCACAG 1260

QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
 DB CTGCTCTGTGTACTCTTCAATCAAAAGACCAACACTTAATGGGCACTGTGCTTATATT 1320
 QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIle 460
 DB TTTGGACTATCTCTGAAATTACTGGAGAGAGCCATATCTATTACTTGACAGCCCTTAATC 1380
 QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPhePropheLys 480
 DB TTCTACCTGGTATTAATCTGACAAAGATGATATACATCAATCAGAGGTTCCCATTTAA 1440
 QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
 DB ACTCTCTCAATGGTTACCTCAATCTTTTACCAACATTTGTTGTTCTTATCAGCAAGTAT 1500
 QY 501 LeuPheGlySerGlyThrLeuProPolysLeuAspValPheAspAlaValAlaArg 520
 DB TATTTGAAAGTGGAACCTTGCTCCAAATTAATGATGTAATTTGATGCTGTGTGCGAAG 1560
 QY 521 HisSerGluGlnAsnMetAspLysThrIleLeuValArgAsnGlnAsnIleLysLeuAsn 540
 DB CACAGTGAAGAGAACATGAGCAAGACCATTTCTAGTCAAAATGAAATATCAATTAAT 1620
 QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
 DB GAACCTTGCACTGTGAAACCTCGGACAGACCTTAACCTCAGTTCAACTTTACCAATTA 1680
 QY 1621 GAACTTGCACTGTGAAACCTCGGACAGACCTTAACCTCAGTTCAACTTTACCAATTA 1680
 QY 561 GluAlaLeuLeuAspValAspSerSerProGlyLysGlySerGlyThrGluAspAsnLeuGln 580
 DB GAGGCCCTCTTGATGTTATTCACGTCGAGGGGCTGTGGAGCTGAGATTAATTTACAA 1740

DB 1681 GAGGCCCTCTTGATGTTATTCACGTCGAGGGGCTGTGGAGCTGAGATTAATTTACAA 1740

RESULT 4
 ADV77918
 ID ADV77918 standard; cDNA; 1743 BP.
 XX
 AC ADV77918;
 XX
 DT 24-FEB-2005 (first entry)
 XX
 DE Mouse high affinity choline transporter (mCHT) cDNA.
 XX
 KW Choline transporter; neuromuscular disorder;
 KW autonomic nervous system disorder; central nervous system disorder;
 KW Parkinsons disease; Huntingtons chorea; genetic disorder;
 KW alzheimers disease; degeneration; neurological disease; schizophrenia;
 KW psychiatric disease; myasthenia gravis; immune disorder; gene therapy;
 KW cms-gen.; nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
 KW muscular-gen.; neuroleptic; DNA purification; gene; ss.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1743
 FT /*tag= a
 FT /product= "Mouse CHT protein"
 XX
 PN US2004248838-A1.
 XX
 PD 09-DEC-2004.
 XX
 PF 01-DEC-2003; 2003US-00724806.
 XX
 PR 23-JUL-2001; 2001US-00911077.
 XX
 PA (UYVA-) UNITV VANDERBILT.
 XX
 PI Blakely RD, Apparsundaram S, Ferguson S;
 XX
 DR WPI: 2005-020586/02.
 DR P-PSDB; ADV77919.
 XX
 PT New human high affinity choline transporter (CHT) cDNA, useful for

PT treating a neuromuscular, autonomic or central nervous system disorder,
PT including Parkinson's disease, Huntington's disease, Alzheimer's,
PT schizophrenia.

PS Example 4: SEQ ID NO 23; 73pp; English.

XX The present invention provides polynucleotides encoding novel high
CC affinity choline transporters (CHTs), methods for their use in screening
CC and therapy. The invention is useful for treating neuromuscular
CC disorders, autonomic or central nervous system disorders such as
CC Parkinsons disease, Huntingtons disease, Alzheimers disease,
CC schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is
CC also useful in gene therapy. The present sequence is mouse high affinity
CC choline transporter (mCHT) cDNA. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?docID=US20040248838.

XX Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6,05e-314	Length:	1743
Score:	2993.00	Matches:	580
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	14	Gaps:	0

US-10-724-806-4 (1-580) x ADV77918 (1-1743)

QY	1	MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu	20
DB	1	ATGCTTTCCAGTGGAAAGAGCTGAGCTATATTCCTTCTTACCTCTTATATTTTCG	60
QY	21	ValGlyIleTPAlaAlaTyrPylsThrLysAsnSerGlyAsnProGluGluArgSerGlu	40
DB	61	GTTGGAATATGGGCTGCATGAGAAAACCAAAACAGCGGACACCGAAGAGCGCATGAA	120
QY	41	AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrAla	60
DB	121	GCCATCATAGTCGGGGCCGTGACATTTGGTTGTTGGTTTTCATGACAGCC	180
QY	61	ThrTPValGlyGlyTyrTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys	80
DB	181	ACCTGGTTGGAGGAGGCTACATCAATGAGACAGACAGAGCTGATGGCCAGGTTGT	240
QY	81	GlyLeuAlaTPAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu	100
DB	241	GGTCTAGCTTGGGCTCATGSCACCATTTGATATTCCTGAGCTAATTTTAGTGTCG	300
QY	101	PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln	120
DB	301	TTTTTTGGAAACCTATGCGTTCCAAAGGATATGTACATATGTAGACCATTCAAACAG	360
QY	121	IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlnMetPhe	140
DB	361	ATCTATGGAAGCGATGGGTGGGCTCTTCATCCCTGCACCTGATGGAGAGATGTC	420
QY	141	TyrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp	160
DB	421	TGGGCTGCAGCAATTTCTCTGCATTTAGGGGCCACATCAGCGTATCTTATGTGGAT	480
QY	161	ValAlaIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly	180
DB	481	GTTGAACATACGCTATGCTCTGCACATTTGCATTTCTTAAACCTTAGGGGTGGG	540
QY	181	LeuTyrSerValAlaIleTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr	200
DB	541	CTCTACTCTGTGGCAATATCTGATGTTGTCAGCTATTTGCAATTTTAAATAGACTGTGG	600
QY	201	IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal	220
DB	601	ATCATGTGTCCTTTTGGCCTGTCACTCTGCAGTACCGACATCGGATTCACAGCTGTG	660

QY	221	HisAlaLysTyrGlnSerProTyrLeuGlyThrIleGlnSerValGluValTyrThrTyrP	240
DB	661	CATGCTAATATACAGAGTCCCTGGCTGGGAACCATTTGAATCATGTTGAGTCTACACCTGG	720
QY	241	LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArg	260
DB	721	CTTGATATTTTCTGTATTTATGCTGGTGGAAATCCATGSCAAGCCTACTCTCCAGAGG	780
QY	261	ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaAlaPheGly	280
DB	781	GTCCTCTTTCATCTCTGACCAACCATCTCAGAGTACGTCTTCTCGGAGCTTTTGGG	840
QY	281	CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyrP	300
DB	841	TGCCGTGGATGGCTCTACCGCCCTATAGCATAGAGCTATTTGGAGTCTCCACACATGG	900
QY	301	AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluAlaAspMetIleLeu	320
DB	901	AACCACTGCTCTACGGGTATCCAGATCCCAAGACTTAAGAGAGACATGATTTCTC	960
QY	321	ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal	340
DB	961	CCGATCGTTCTGCAAGTACCTCGCCCTGTATACATCTCTTCTTGGGCTGTGCTGTT	1020
QY	341	SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla	360
DB	1021	TCAGCTGCTGATGCTCTCCTCAGCTGACCTGCTCATCTGCTGGCCAGTTCATATGTTGCT	1080
QY	361	ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrVal	380
DB	1081	CGGAATATCTACACACTTCTTCCAGACAAATGATACAGACAGAAATTTGTGGGCTC	1140
QY	381	MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys	400
DB	1141	ATGAGATCAGCTGCTGTTGTTGGAGCATCTGCACAGCATGCTTGTGTCAGAAAG	1200
QY	401	ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln	420
DB	1201	ACTGTGATGGGCTCTGGTACTGCTGACCTGCTGCTGCTCATCATATCTTCCACAG	1260
QY	421	LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle	440
DB	1261	CTGCTCTGTATCTTTATCAAAAGAACCAACACTTATGGGCGAGTTGCTGTATATTT	1320
QY	441	PheGlyLeuPheLeuArgIleThrGlyGlyGlnProTyrLeuTyrLeuGlnProLeuIle	460
DB	1321	TTTGACATATTCTGAGATTTACTGAGAGAGCCATATCTATACCTTGACGCCCTTAATC	1380
QY	461	PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys	480
DB	1381	TTCTACCTCGTTTATTTACTCTGACAAAGATGGTATATCAACCAAGGTTCCCATTTTAA	1440
QY	481	ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaIysTyr	500
DB	1441	ACTCTCTGCATGTTAATCTCATTTCTTACCAACATTTGTGTTCTTATCTAGCCAAAGT	1500
QY	501	LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg	520
DB	1501	CTATTGAAAGTGAACCTGCTCCAAATATAGATGATTTGATGCTGTTCTCCCAAGG	1560
QY	521	HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn	540
DB	1561	CACAGTGAAGAGAACATGAGAACCAATTTCTAGTCAGAAAAGAAAATATCAAAATTAAT	1620
QY	541	GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys	560
DB	1621	GAACTTGACCTGTGAAACCTCGGCAAGCCTTAACCTCAGTTCACCAATTAAAG	1680
QY	561	GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln	580
DB	1681	GAGGCCCTCTCTTGAATGTTGATTCAGTCCGAGGGGCTGTGGGACATGAAGATTAATTACA	1740

RESULT 5
AAFB1713
ID AAFB1713 standard; cDNA; 1743 BP.
AC AAFB1713;
XX
DT 01-JUN-2001 (first entry)
XX
DE Mouse high affinity choline transporter protein encoding cDNA.
XX
KM High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
XX ss.
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT 1.1743
FT CDS /product= "high affinity choline transporter"
XX
PN MO200116315-A1.
XX
PD 08-MAR-2001.
XX
PF 18-AUG-2000; 2000MO-JP005545.
XX
PR 27-AUG-1999; 99JP-00240642.
XX 27-DEC-1999; 99JP-00368991.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Haga T, Okuda T;
XX
DR MPI; 2001-226688/23.
XX P-PSDB; AAB74666.
XX
DR New rat and human spinal cord high affinity choline transporters, useful
XX in diagnosis of Alzheimer's disease and screening promoters as drugs for
XX treating Alzheimer's disease.
XX
PS Claim 12; Page 78-82; 90pp; Japanese.
XX
XX The present sequence encodes a mouse (Mus musculus) high affinity choline
XX transporter protein designated cho-1. The cho-1 protein has nootropic and
XX neuroprotective activities. The cho-1 polynucleotide and protein can be
XX used for the diagnosis of diseases related to the expression of cho-1 by
XX comparing the cho-1 polynucleotide sequence in a sample to that of a
XX control. Drug compositions containing the cho-1 protein or expression
XX promoters or inhibitors of cho-1 are useful for treating disorders
XX characterised by abnormal levels of cho-1, such as Alzheimer's disease
XX
SQ Sequence 1743 BP; 407 A; 410 C; 409 G; 517 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3.12e-311 Length: 1743
Score: 2968.00 Matches: 576
Percent Similarity: 99.5% Conservative: 1
Best Local Similarity: 99.3% Mismatches: 3
Query Match: 99.2% Indels: 0
Gaps: 0
US-10-724-806-4 (1-580) x AAFB1713 (1-1743)
QY 1 MetProPhHisValGluGlyLeuValAlaIleIleuPheTyrlleuIleuIlePheIeu 20
Db 1 AATCTCTTCACAGTAAGAGCTGTAGCTATTATCTCTTCTTCACTCTTATATTCTG 60
QY 21 ValGlyIleTrrPalaIaIaTrrPlyStnIlyAsnSerGlyAsnProGluGluAtsSerGlu 40
Db 61 GTTGGAATATGGCTGATGAGAAACCAAAACAGGGCAACCAAGAGAGCAAGTGA 120
QY 41 AlaIleIleValaIlyGlyArgAspIleGlyLeuIleuValaIlyGlyPheThrMetTrnAla 60
|||||

Db 121 GCCATCATAGTCGGGGCCGTGACATTGGTTGTGGTTGGTTTACATGACAGCC 180
QY 61 ThrTrpValaIlyGlyTyrlleuAsnGlyTrnAlaGluAlaValTyrlGlyProGlyCys 80
Db 181 ACCTGGTTGGAGAGGCTATCATCATGAGACAGCAAGCACTGATATGGCCAGTTGT 240
QY 81 GlyLeuAlaIaTrrPalaHisAlaProIleGlyTyrlSerleuSerleuIleuGlyIleu 100
Db 241 GGTCTAGCTTGGGCTCAGGACCACTTGATATTTCTGAGTCTAATTTAGTGGTCTG 300
QY 101 PhePheAlaIyProMetArgSerIlyGlyTyrlValThrMetLeuAspProPheIySgl 120
Db 301 TTTTTCGAAACCTATGGCTTCAGAGGATATGACTATGTAAACCATTTTCAACAG 360
QY 121 IleTyrlGlyLeuArgMetGlyIlyLeuLeuPheIleProAlaIleuMetGlyIleuPhe 140
Db 361 ATCTATGAAAGGCAATGGGTGGGCTGCTCTTCATCTGCACTGATGGAGAGATGTT 420
QY 141 TrpAlaAlaAlaIlePheSerAlaIleuGlyAlaThrIleSerValIleIleAspValasp 160
Db 421 TGGGCTGCAAGCAATTTCTCTGCAATTAGGGGCCACCATCAGCGTATCATTTGAT 480
QY 161 ValAsnIleSerValIleValSerAlaIleuIleAlaIleuTyrlThrIleuValaIlyGly 180
Db 481 GTGAACATATCGGTCAATGTCTCTGCACTCATTCATTCATTTATACCCCTAGGGGTGG 540
QY 181 IeuTyrlSerValaIaIaTrrPalaIaIaValaIleuPheCysIlePheIleGlyIleuTrp 200
Db 541 CTCTACTCTGTGGCATATCTGATGTGTGTCAGCTATTCGCAATTTTATAGACGTGG 600
QY 201 IleSerValProPheAlaIleuSerHisProAlaAlaIThrAspIleGlyPheThrAlaVal 220
Db 601 ATCAGTGTCCCTTTTGGCTGTACATCCTGCACTGACGACATCGGATTCACAGCTGTG 660
QY 221 HisAlaIyTyrlGlnSerProTrpLeuGlyThrIleGluSerValaIlyTyrlTrp 240
Db 661 CATGCTAAATATACAGAGTCCCTGCTGGAGAACATTAATCACTTAAGCTTACACCTGG 720
QY 241 IeuAspAsnPheIleuIleuMetLeuGlyIleProTrpGlnAlaTyrlPheGlnArg 260
Db 721 CTGATTAATTTCTGTATATGATGTGGGTGGAGATCCATGGAGACTTATCCAGAGG 780
QY 261 ValIleuSerSerSerSerAlaThrTyrlaIleValIleuSerPheIleuAlaIaPheGly 280
Db 781 GTCTCTCTTCACTCCAGCACTATGCTCAGTACTGTCTTCCGGGAGCTTTTGGG 840
QY 281 CysLeuValMetAlaIleuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db 841 TGCCTGTATGCTCTTACCCGCAATATGATCATAGAGCTTATGAGACTTCCACAGACTGG 900
QY 301 AsnGlnThrAlaIyTyrlTyrlProAspProIlySThrIleGluGlnAlaAspMetIleu 320
Db 901 AACCAAGACTGCTTACGGATATCCAGATCCCAAGACTAAGAGAGAACAGATGATTTCTC 960
QY 321 ProIleValIleuGlnTyrlleuCysProValTyrlleSerPhePheGlyIleuGlyAlaVal 340
Db 961 CCGATGTCTTGAAGTACTTGCCTCTGTGATCATCTCTTCTTGGGCTGGTGTGCTT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleuSerAlaSerMetPheAla 360
Db 1021 TCAGCTGCTGTCATGCTCTCAGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 361 ArgAsnIleTyrlGlnIleuSerPheAspGlnAsnAlaSerAspIyGlyIleValIleTrpVal 380
Db 1081 CGGAATATTCACAGCTTTCTTCTTCAAGCAAAATGATCAGACAGAAATGTGTGGTCT 1140
QY 381 MetArgIleThrValIleuValPheGlyAlaSerAlaThrAlaMetAlaIleuIleuTrpLys 400
Db 1141 ATGAGATCATCTGTGCTTGTGTCGAGCATCTGCAACAGCATGTGCTTGTGTCAGAAAG 1200
QY 401 ThrValTyrlGlyIleuTrpTyrlleuSerSerAspLeuValTyrlIleIlePheProGln 420
Db 1201 ACTGTATATGGGCTGTGATCTGAGCTGACCTGTGTCTTACATCATCTTCCACAG 1260
|||||

OY 421 LeuLeuGysValLeuPheLeileysGlyThrAsnThrTYRGLyAlaValAlaGlyTyrIle 440
DB 1261 CTGCTCTGTGTAAGTCTTCAATCAAGAAACCAACTTATGGGCGAGTGTGCTTATATTT
OY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGlyProTYRLeuTYRLeuGlnProLeuIle 460
DB 1321 TTTGGACTATTTCTTGAAGATTTCTGGAGAGGACCAATCTATCTTACTTGCAGCCCTTAATC 1380
OY 461 PheTYRProGlyTYRTYRserAspIysAsnGlyIleTYRAsnGlnArgPheProPheIys 480
DB 1381 TTCTACCTGGGTTATTAATCTGACAGAAATGGTATATACAACTACAGAGGTTCCCATTTAA 1440
OY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleGysValSerTYRLeuAlaIysTYR 500
DB 1441 ACTCTCTCATGTGTAACCTCATCTTCTTACCAACATTTGGTGTCTTATCTTACCAAGTAT 1500
OY 501 LeuPheGlySerGlyTYRThrLeuProProlYsLeuAspValPheAspAlaValAlaArg 520
DB 1501 CTATTTGAAAGTGGAACTTGGCTCCCAAAATTAGATTTTGATGCTGTGTTCGCAAG 1560
OY 521 HisSerGlyGluAsnMetAspIysThrIleLeuValArgAsnGluAsnIleYsLeuAsn 540
DB 1561 CACAGTGAAGAGAAATGAGAACCAATCTTACGTACAGAAATGAAATCAATCAATTAAT 1620
OY 541 GlnLeuAlaProValIysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIys 560
DB 1621 GAACTTGACCTGTGAAACCTTGCGAGAGCCCTTACCCCTCATCTTCAACATTAAG 1680
OY 561 GlnAlaLeuLeuAspValAspSerSerProGlyGlySerGlyTYRThrGluAspAsnLeuGln 580
DB 1681 GAGGCGCTCTTGAAGTGTATTCAGTCCGAGGGGTCTGGAGCTGAGATTAATCTTACAA 1740

RESULT 6

AAD02457
ID AAD02457 standard; cDNA; 4938 BP.

AC AAD02457;

DT 24-APR-2001 (first entry)

DE Mouse P4P6B1 OMA (obese mice adipocyte) protein encoding cDNA.

KM Mouse; OMA protein; obese mice adipocyte; P4P6B1;

KW fuel metabolism disorder; therapy; obesity; diabetes; gene therapy;

KM anorectic; antidiabetic; ss.

OS Mus sp.

XX Key Location/Qualifiers

FT CDS 247..1989 /*tag= a

FT misc_feature /product= "Mouse P4P6B1 OMA protein"

FT /*tag= b 988..1342

FT /note= "Portion of original 450 bp PCR fragment"

XX WO200078950-A2.

XX 28-DEC-2000.

XX 13-JUN-2000; 2000MO-US016217.

XX 22-JUN-1999; 99US-014151SP.

XX (AMYL-) AMYLIN PHARM INC.

XX Sierzega M, Albrandt K;

XX WPI, 2001-112322/12.

XX P-PSDB; AAAT72388.

PT Novel obese mice adipocyte polypeptides useful in diagnosis and treatment

PT of disorders of fuel metabolism such as obesity or diabetes.

XX Claim 2; Fig 3; 83pp; English.

XX The present sequence is mouse P4P6B1 cDNA which encodes OMA (obese mice
XX adipocyte) protein. The P4P6B1 fragment was generated by RNA
XX fingerprinting using random primers P4 and P6. OMA is used as a
XX diagnostic reagent for diagnosing a disorder of fuel metabolism in an
XX underweight or an overweight individual, by detecting the transcription
XX level of a gene encoding OMA, which is induced or repressed in an
XX individual by a factor such as genetic obesity, fasting and refeeding of
XX a fasted individual. OMA is useful in the generation of antibodies, for
XX use in pharmaceutical compositions and for studying DNA/protein
XX interactions. Nucleic acids encoding OMA are involved in gene therapy. An
XX inhibitor of OMA or an antisense oligonucleotide that inhibits expression
XX of OMA are useful for treating disorders of fuel metabolism such as
XX obesity or diabetes

XX Sequence 4938 BP; 1436 A; 1012 C; 976 G; 1514 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.98e-310 Length: 4938
Score: 2967.00 Matches: 576
Percent Similarity: 99.54 Conservative: 1
Best Local Similarity: 99.34 Mismatches: 3
Query Match: 99.18 Indels: 0
DB: Gaps: 0

US-10-724-806-4 (1-580) x AAD02457 (1-4938)

OY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTYRLeuLeuIlePheLeu 20

DB 247 ATGCTTTCCAGTGAAGAGAGCTGTACCTATTAATCTCTTCTTACCTCTTATCTTCTG 306

OY 21 ValGlyIleTrrPAlaAlaTrrPlySerThrIysAsnSerGlyAsnProGluGluArgSerGlu 40

DB 307 GTTGGAAATATGGGCGTGCATGGAAGAAACCAAAACAGCGGCAACCCAGAGAGCGCATGA 366

OY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60

DB 367 GCCATCATAGTCGGGGCGGTGACATTTGTTGTTGGTGGTTTATCCATGACAGCC 426

OY 61 ThrTrrValGlyGlyTYRThrIleAsnGlyThrAsnGlyAlaValTYRGLYProGlyCys 80

DB 427 ACCTGGGTGAGAGGCTACATCATAGAGACAGAGAGCATGATATGCGCTAGGTGT 486

OY 81 GlyLeuAlaTrrPAlaHisAlaProIleGlyTYRserLeuSerLeuIleLeuGlyGlyLeu 100

DB 487 GGTCTAGCTGGGCTCAGGCAACCCATTGGATATTTCTCGAGCTTAAATTTAGTGTCTG 546

OY 101 PhePheAlaIysProMetArgSerIysGlyTYRValThrMetLeuAspProPheIysGln 120

DB 547 TTTTGGGCAACCTATGCTGTCCAAAGGATATGACTATGTTAGACCCATTTCACAG 606

OY 121 IleTYRGLYsArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140

DB 607 ATCTATGAAAGCGATGGTGGCTGCTCTTCACTCCCTGACCTATGGAGAGATGTC 666

OY 141 TrrPAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160

DB 667 TGGGCTGAGCAATTTTCTGTCATTAGGGGCAACATCAGGTGATCATTTGATGTGAT 726

OY 161 ValAsnIleSerValIleValIserAlaLeuIleAlaIleLeuTYRThrLeuValGlyGly 180

DB 727 GTGAACATATGCTCATTTCTCTGACCTCATTTGCACTTCTTATATACCTATGTTGG 786

OY 181 LeuTYRserValAlaTYRThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrrP 200

DB 787 CTCTACTCTGGCGCATATCTGATGTGTCCAGCTATTTCTGATTTTATAGACTGTGG 846

OY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220

DB 847 ATCACTGTCCCTTTGGCCCTGTGCATCATCTGCAAGTACGACATCGGATTCACACTGTG 906


```

QY 221 HisAlaIyTyGlnSerProFpleuGlyThrIleGluSerValGluValTyThrTrp 240
DB 907 CATGCTAAATACCAAGATCCTGGCTGGGAACCATTAATCACTTAAGCTTACACCTGG 966
QY 241 LeuAspAsnPhelLeuLeuMetLeuGlyIleProTyrGlnAlaTyRphneGlnArg 260
DB 967 CTTGATTAATTTCTGTTATGATGCTGGGGATGCCATGCAAGCCTTACCTCCAGAG 1026
QY 261 ValLeuSerSerSerSerAlaThrTyRAlaGlnValLeuSerPhelLeuAlaIlePheGly 280
DB 1027 GTTCCTCTCTTCATCCCTCAAGCCACCTATGCTCAGATCTGCTTCTTGAGCTTTGGG 1086
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
DB 1087 TGCTTGATATGAGCTTACCCGCAATATGATATGAGCTTATGGAGCTTCCACAGACTGG 1146
QY 301 AsnGlnThrAlaTyGlyTyRProAspProLysThrLysGluGluAlaAspMetIleLeu 320
DB 1147 AACCAAGACTGCGTACGGGATCCAGATCCCAAGACTTAAGAGAGACAGCATGATTTCTC 1206
QY 321 ProIleValLeuGlnIleTyRLeuCysProValTyRLeuSerPheGlyLeuGlyAlaVal 340
DB 1207 CCGATGCTTCGACAGTACCTGCTCCCTGTGTACATCTCTTCTTGGGCTTGGCTGTGT 1266
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1267 TCAAGCTGCTGTATGCTTCAAGCTGCTGCTCATCTGTGCGAGCTTCTATGTTTGGT 1326
QY 361 ArgAsnIleTyGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyRVal 380
DB 1327 CGGAATATCTACAGACTTTCCTTCCAGCAAAATGATCAGACAGAAATGTGTGGGTC 1386
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
DB 1387 ATGAGAGATCACTGTGTTGTGTTCGAGATTCGCAACGCCATGTGCTTGCAGCAAG 1446
QY 401 ThrValTyRGlYLeuThrTyRLeuSerSerAspLeuValTyRleIleIlePheProGln 420
DB 1447 ACTGTATAGGGCTCTGGTACTGAGCTCTGACCTGTCTACATCATCTTCCACAG 1506
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyRGlYAlaValAlaGlyTyRle 440
DB 1507 CAGCTCTGTGATCTTTCATCAAAAGAACCAACATTATGGGCGAGTGTGTTATATT 1566
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyRLeuTyRLeuGlnProLeuIle 460
DB 1567 TTTGGACTATCTCGAATTAATCTGAGAGAGACCATATCTATATCTTGACGCCCTTAATC 1626
QY 461 PheTyRProGlyTyRLeuSerAspLysAsnGlyIleTyRAsnGlnArgPheProPheLys 480
DB 1627 TTCTACCCCTGGTATTAATCTGACAAAGATGATATACATACAGAGTTCCTCATTTAA 1686
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyRLeuAlaLysTyR 500
DB 1687 ACTCTCTCCATGTTTACCTCATCTTACCAACATTGTGTTCTTATATCAGCAAGTAT 1746
QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaIleArg 520
DB 1747 CTAATTTGAAGTGAAGCTTGGCTCCCAAAATTAGATGATTTGATGTTGTGGCAAGC 1806
QY 521 HisSerGluGlnAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
DB 1807 CACAGTGAAGAGACATGAGCAAGACCATTTCTAGTCAAGAAATGAATATCAAAATTAAT 1866
QY 541 GluLeuAlaProValLysProArgIleSerLeuThrLeuSerSerThrPheThrAsnLys 560
DB 1867 GAATCTGCACCTGTGAACCTCGGAGAGGCTTAACCTTCAGTTCAACTTTCAACAATAAG 1926
QY 561 GluAlaLeuLeuAspValaAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
DB 1927 GAGGCCCTCTTATGATTTGATTCAGTCCGAGAGGGGTCTGGAGCTGAAGATTAATCAAA 1986

```

```

RESULT 7
AAFB1711
ID AAFB1711 standard; cDNA; 1743 BP.
XX
AC AAFB1711;
XX
DT 01-JUN-2001 (first entry)
XX
DE Rat high affinity choline transporter protein encoding cDNA.
XX
KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
XX ss.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 1..1743
FT /tag= a
FT /product= "high affinity choline transporter"
XX
PN MO200116315-A1.
XX
PD 08-MAR-2001.
XX
PF 18-AUG-2000; 2000MO-JP005545.
XX
PR 27-AUG-1999; 99JP-00240642.
PR 27-DEC-1999; 99JP-00368991.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Haga T, Okuda T;
XX
DR WPI; 2001-226688/23.
XX
DR P-PSDB; AAB74664.
XX
PT New rat and human spinal cord high affinity choline transporters, useful
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
PT treating Alzheimer's disease.
XX
PS Claim 6; Page 64-68; 90pp; Japanese.
XX
XX The present sequence encodes a rat (Rattus norvegicus) high affinity
XX choline transporter protein designated cho-1. The cho-1 protein has
XX nootropic and neuroprotective activities. The cho-1 polynucleotide and
XX protein can be used for the diagnosis of diseases related to the
XX expression of cho-1 by comparing the cho-1 polynucleotide sequence in a
XX sample to that of a control. Drug compositions containing the cho-1
XX protein or expression promoters or inhibitors of cho-1 are useful for
XX treating disorders characterised by abnormal levels of cho-1, such as
XX Alzheimer's disease
XX
SQ Sequence 1743 BP; 414 A; 402 C; 404 G; 523 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 6,266-310 Length: 1743
Score: 2956.00 Matches: 572
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 98.6% Mismatches: 4
Query Match: 98.8% Indels: 0
DB: 4 Gaps: 0
US-10-724-806-4 (1-580) x AAFB1711 (1-1743)
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyRLeuLeuIlePheLeu 20
DB 1 ATGCCCTTCCATGTAAAGACTAGTAGCATTAATCTGTCTTACCTTCTATATTTCTG 60
QY 21 ValGlyIleTyRAlaIleTyRProLysAsnSerGlyAsnProGluGluArgSerGlu 40
DB 61 GTTGGAAATATGGCTCATATGAGAAACCAAAACAGCGGTATATCCAGAAACCAAGCGAA 120
QY 41 AlaIleIleValGlyIleArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60

```


XX Novel isolated polynucleotide encoding human or mouse high affinity
PT choline transporter polypeptide, useful in gene therapy to increase
PT cholinergic function in a cell of a patient suffering from Alzheimer's
PT disease.

XX Example 1; SEQ ID NO 5; 74bp; English.

XX The present invention relates to the isolation of polynucleotide
CC sequences encoding human and mouse high-affinity choline transporter
CC (hCHT and mCHT respectively), and the proteins they encode. The gene
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
CC encoding hCHT is useful for expressing hCHT recombinantly. The hCHT
CC polynucleotide sequence when delivered to a cell, increases cholinergic
CC function in the cell that is in a patient having Parkinson's disease.
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
CC myasthenia gravis. The hCHT antibody is useful for controlling
CC transporter CHT proteins to the brain, and for treating the above
CC mentioned diseases. The antibody is also useful for diagnosing the above
CC mentioned disorders and to detect the influence of cholinergic
CC signalling. The present sequence encodes rat CHT (rCHT). Note: The
CC sequence data for this patent was obtained in electronic format directly
CC from the USPTO web site at seqdata.uspto.gov.

XX Sequence 4904 BP; 1447 A; 991 C; 939 G; 1527 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,06e-309 Length: 4904
Score: 2956.00 Matches: 572
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 98.6% Mismatches: 4
Query Match: 10 Indels: 0
DB: Gaps: 0

US-10-724-806-4 (1-580) x ADD50642 (1-4904)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrlleuLeuIlePheLeu 20
DB 224 ATGCTTTCATGTAG 283
QY 21 ValGlyIleThrAlaAlaTyrIleThrIleAsnSerGlyAsnProGluIleArgSerGlu 40
DB 284 GTTGGAAATATGGGCTGCATGAGAAACCAAAACAGCGGTATGCAGAGAACCCAGCGAA 343
QY 41 AlaIleIleValGlyIleArgAspIleGlyLeuLeuValGlyIlePheThrMetThrAla 60
DB 344 GCCATCATGTGGGGGCGAG 403
QY 61 ThrTTPValGlyIleGlyIleThrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
DB 404 ACCTGGGTGGAGTGT 463
QY 81 GlyLeuAlaIleThrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyIleu 100
DB 464 GGTCCTAGCTTGGGCTCAGGACCCCATTTGATATCTCTGAGTCTGATTTAGGTGGCTG 523
QY 101 PhePheAlaIlePhePheMetIleArgSerIleGlyTyrValThrMetLeuAspProPheGln 120
DB 524 TTTTGTGCAAAACCTATGGCTTCAGAGGATATGTAATGTAATGTAATGTAATGTAATG 583
QY 121 IleTyrGlyIleArgMetGlyIleLeuPheIleProAlaLeuMetGlyIleMetPhe 140
DB 584 ATCTATGAGAAAGGAGTTC 643
QY 141 TTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB 644 TGGGCTGCAGCAATTTCTCTGCAATTAAGGAGGCTAATGCTAATGCTAATGCTAATG 703
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyIle 180
DB 704 GTGAAATATTCGGTCTATGCTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 763
QY 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200

DB 764 CTTACTCTGTGGCATATCTATGTTGTATGAGCTATTCGATTTTATAGATTTGTGG 823
QY 201 IlesSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyIlePheThrAlaVal 220
DB 824 ATCAGGTCCCAATTTGGCTGCTGCATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 883
QY 221 HisAlaIleTyrGlnSerProIlePheGlyIleIleGlySerValGluValTyrThrTyr 240
DB 884 CATGCTAATATCCAGAGTCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 943
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleProTyrGlnAlaTyrPheGlnArg 260
DB 944 CTGTAATATTTCTGTTGTATGCTGGGGTGAATACATGAGCAACCTTATCCAGAGG 1003
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 1004 GTCTCTCTTCATCCGACAG 1063
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300
DB 1064 TGCTGTATGAGCTTACAGAGCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1123
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProIlePheGlyIleGlyIleLeu 320
DB 1124 AACCAATGATATGAGTTCAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
QY 321 ProIleValLeuGlnTyrIleuCysProValTyrIleSerPhePheGlyIleGlyAlaVal 340
DB 1184 CGATTTGCTTCACAGATCTGAG 1243
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1244 TCTGCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1303
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspIleGlyIleValTTPVal 380
DB 1304 CGGAATATCTACAGAGTTCCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1363
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400
DB 1364 ATGAGAGATACGTGTTGTGTTGGAGATCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1423
QY 401 ThrValTyrGlyIleuTyrIleuSerSerAspLeuValTyrIleIleIlePheProGln 420
DB 1424 ACTGTATAGGGCTCGTACCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1483
QY 421 LeuLeuCysValLeuPheIleIleGlyIleThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
DB 1484 CTGCTCTGTGATCTTCTTCAAG 1543
QY 441 PheGlyIleuPheLeuArgIleThrGlyIleGlyIleuProTyrIleuIleuProLeuIle 460
DB 1544 TTTGACTTTTCTGAGATTAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1603
QY 461 PheTyrProGlyTyrTyrIleSerAspIleGlyIleTyrAsnGlnArgPheProPheIle 480
DB 1604 TTCTACCTGCTTATTAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1663
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaIleTyr 500
DB 1664 ACTCTCTCATGATTCATCTTATCTTATCAACATTTGTTGTTCTATGAGCAAGATAT 1723
QY 501 LeuPheGlnSerGlyIleuPheProIleuPheAspValPheAspAlaValAlaIleArg 520
DB 1724 CTATTTGAAAGGAACTTTGCTCCAAATTAAGATATTTGATCTGTTGCTCAAGG 1783
QY 521 HisSerGlyIleuAsnMetAspIleThrIleLeuValArgAsnGlnIleuIleuIleuAsn 540
DB 1784 CACAGTGAAG 1843
QY 541 GluLeuAlaProValIlePheArgGlnSerLeuThrLeuSerSerThrPheThrAsnIle 560

Db 1844 GAACCTGCACCTGTAAGCCTCGACAGACCTTAACCTTCACTTACCAATATA 1903

Qy 561 G|u|l|a|l|e|u|a|s|p|a|s|e|r|e|p|r|o|g|l|y|s|e|r|g|l|y|t|h|r|g|l|u|a|s|p|a|s|e|r|e|u|g|l|n| 580

Db 1904 GAGGCTCTCCTTGATGATTGATTCAGAGGAGATCTGGGACTGGAAGATTAATTACA 1963

RESULT 9

ID ADV77900 standard; cDNA; 4904 BP.

XX ADV77900;

AC ADV77900;

XX 24-FEB-2005 (first entry)

DT Rat high affinity choline transporter (rCHT1) cDNA.

DE Choline transporter; neuromuscular disorder;

XX autonomic nervous system disorder; central nervous system disorder;

KM Parkinsons disease; Huntingtons chorea; genetic disorder;

KM Alzheimers disease; degeneration; neurological disorder; schizophrenia;

KM psychiatric disease; myasthenia gravis; immune disorder; gene therapy;

KM cns-gen.; neuroprotective; antiparkinsonian; anticonvulsant;

XX muscular-gen.; neuroleptic; DNA purification; gene; ss.

OS Rattus norvegicus.

XX Key Location/Qualifiers

FT CDS 224..1966

FT /*tag= a

FT /product= "Rat CHT protein"

XX

PN US2004248838-A1.

XX

PD 09-DEC-2004.

XX

PF 01-DEC-2003; 2003US-00724806.

XX

PR 23-JUL-2001; 2001US-00911077.

XX

PA (UYVA-) UNIV VANDERBILT.

XX

PI Blakely RD, Apparsundaram S, Ferguson S;

XX WPI. 2005-020586/02.

DR P-PSDB; ADV77901.

XX

DR GENBANK; AB030947.

XX

PT New human high affinity choline transporter (CHT) cDNA, useful for

PT treating a neuromuscular, autonomic or central nervous system disorder,

PT including Parkinson's disease, Huntington's disease, Alzheimer's,

PT schizophrenia.

XX

PS Example 1; SEQ ID NO 5; 73bp; English.

XX

CC The present invention provides polynucleotides encoding novel high

CC affinity choline transporters (CHTs), methods for their use in screening

CC and therapy. The invention is useful for treating neuromuscular

CC disorders, autonomic or central nervous system disorders such as

CC Parkinsons disease, Huntingtons disease, Alzheimers disease,

CC schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is

CC also useful in gene therapy. The present sequence is rat high affinity

CC choline transporter (rCHT1) cDNA. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from the US patent office at

CC segdata.uspto.gov/sequence.html?DocID=US20040248838.

XX

CC Sequence 4904 BP; 1447 A; 991 C; 939 G; 1527 T; 0 U; 0 Other;

SO

Alignment Scores:

Pred. No.: 3.06e-309 Length: 4904

Score: 2956.00 Matches: 572

Percent Similarity: 99.3% Conservative: 4

Best Local Similarity: 98.6% Mismatches: 4

Query Match: 98.8% Indels: 0

DB: 14 Gaps: 0

US-10-724-806-4 (1-580) x ADV77900 (1-4904)

Qy 1 MetProPheHsValGluGlyLeuValAlaIleIleLeuPheTyrLeuIlePheLeu 20

Db 224 ATGCTTTCATGTAAGAAAGACTAGACGATATCTGTTCTTACCTTTATATTTCTG 283

Qy 21 ValGlyIleTrrPalAlaIleTrrLysThrLysAsnSerGlyAsnProGluLysSerGlu 40

Db 284 GTTGAATATGGGCTGCATGGAACCAAAACAGCGGTAATGCAAGAAACGACGCA 343

Qy 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrAla 60

Db 344 GCCATCATGATGGGGGCCGAGACATTTGTTGTGGTGGTGTATGACATGACAGCC 403

Qy 61 ThrTrrValGlyGlyGlyTrrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80

Db 404 ACCTGGTTGGAGGGGTTTACATCAACGGACGCTGAAGAGATTATGGCCAGGTTGT 463

Qy 81 GlyLeuAlaTrrPalAlaIleAlaProIleGlyTyrSerLeuSerLeuIleGlyLeu 100

Db 464 GGTCTAGCTTGGGCTCAGGACCCATGGATATTTCTGAGTCTGATTTTAGTGCCCTG 523

Qy 101 PhePheAlaLysProMetArgSerLysGlyTyrValTrrMetLeuAspProPheLysGln 120

Db 524 TTTTTCGAAAACCTATGCGTTCCAAAGGATATGTGACTATGTATGACCCGTTTCAACAG 583

Qy 121 IleTrrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140

Db 584 ATCTATGAAACCGATGGGTGGGCTGCTTTCATCCCTGCACGTATGGAGAGATGTTT 643

Qy 141 TrrPalAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValLsp 160

Db 644 TGGGCTGCAGCAATTTTCTGCAATTAGGGGCTACATCAGGTAATCATGATGTGAT 703

Qy 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180

Db 704 GTGAACATATGGGTATGTCCTCGCACTATGGCATTTTATACCTCTGGGAGGG 763

Qy 181 LeuTyrSerValAlaTyrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrr 200

Db 764 CTCTACTCTGGGCAATATCATGATGTTGTACAGCTATTTCTGATTTTATAGATTGG 823

Qy 201 IleSerValProPheAlaLeuSerHisProAlaValTrrAspIleGlyPheThrAlaVal 220

Db 824 ATCAGTGTCCCATTTGGCCCTGTCAATCTGCAATGACGACATGGAATTCAGTCTG 883

Qy 221 HisAlaLysTrrGlnSerProTrrLeuGlyThrIleGlySerValGluValTyrThrTrr 240

Db 884 CATGCTAAATACCAAGTCCCTGGCTGGAAACCATGTAATCATGTAAGTCTACACCTGG 943

Qy 241 LeuAspAsnPheLeuLeuMetLeuGlyIleProTrrGlnAlaTyrPheGlnArg 260

Db 944 CTGTATATTTTCTGTTGTGATGCTGGGTGGAATACATGGCAAGCTACTTCCAGAGG 1003

Qy 261 ValLeuSerSerSerSerAlaThrTrrAlaGlnValLeuSerPheLeuAlaIlePheGly 280

Db 1004 GTCCTCTTTCATCTGTCAGGACCATATCTCAGGGCTGTCTCTTCTGAGCTTTTGGG 1063

Qy 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrr 300

Db 1064 TCCCTGGTGAATGGCTTCACGAGCCATTTGGCATTTGGGGCCATTTGGACCTCCACAACCTGG 1123

Qy 301 AsnGlnThrAlaTyrGlyTrrProAspProLysThrLysGluGluAlaAspMetIleLeu 320

Db 1124 AACCAATGCTATATGGGTTTCCAGATCCCAAGCAAGAGAGAGACATGATTTCTC 1183

Qy 321 ProIleValLeuGlnTrrLeuCysProValTrrIleSerPhePheGlyLeuGlyAlaVal 340

Db 1184 CCGATGTTTCTACAGTACTGCTGACCTGTGATCATTTCTTCTTGGGCTTGCTGTT 1243

```

OY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1244 TCTGCTGCTGTCATGTCCTCGGCTGACATCATCTCATCAGCAAGTTCATGTTGCT 1303
OY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspGlnIleValTTPVal 380
DB 1304 CGGAATATCTTACACACTTCTTCCAGACAAAATGATCAGACAGAAATTTGTGGGTC 1363
OY 381 MetArgIleThrValLeuValPheGlnAlaSerAlaThrAlaMetAlaLeuMetThrLys 400
DB 1364 ATGAGATCATCTGTTGTTGTTGTTGAGCATCTGCACAGCCATGGCTTGCCTCAGAG 1423
OY 401 ThrValTyrGlnLeuTyrTyrLeuSerSerAspLeuValTyrIleIleLeuPheProGln 420
DB 1424 ACTGTGTATGGGCTCTGGTACCTGAGCTGACCTGTCTATCATCATCTTCCACAG 1483
OY 421 LeuLeuCyValLeuPheIleIleGlyThrAsnThrTyrGlnValAlaIleGlyTyrIle 440
DB 1484 CTGCTCTGTGTACTCTTCAACAAAGAACCAACATTATGGGGCACTTCTGTATATAT 1543
OY 441 PheGlnLeuPheLeuArgIleThrGlnGlnGluProTyrLeuTyrLeuGlnProLeuIle 460
DB 1544 TTGGACTTTTCTCGAGATTCACGAGAGAGCCATATCTTACTTGCAGCCCTTAATC 1603
OY 461 PheTyrProGlnTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
DB 1604 TTCTACCCGTTATTAACCTGACAAAGATGGATATACATCAGAGGTTCCCATTTAA 1663
OY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCyValSerTyrLeuAlaLysTyr 500
DB 1664 ACTCTCTCCATGTTAACCCTCATCTTTTACCAACATTGTTGTTCCATCTAGCAAGTAT 1723
OY 501 LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaIleArg 520
DB 1724 CTATTTGAAAGTGAACCTTGCTCCAAATTAAGATATATTTGATGCTGTGTCTAAG 1783
OY 521 HisSerGlnGluAspMetAspLysThrIleLeuValAlaArgAsnGluAsnIleLysLeuAsn 540
DB 1784 CACAGTGAAGAGAACATGACACAAACCATTTAGTACGAAATGAAAACATCAATTAAT 1843
OY 541 GlnLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
DB 1844 GAACCTTGACCTGTAAAGCCTGCACAGAGCCTAACCTCAGTTCACTTCCACATTA 1903
OY 561 GlnAlaLeuLeuAspValAspSerSerProGlnGlySerGlyThrGlnAspAsnLeuGln 580
DB 1904 GAGGCTCTCCTTGATGTTGATTCACAGTCCAGAGGGATCTGGGCTGAAGATTA 1963
RESULT 10
AAF81712
ID AAF81712 standard; cDNA, 1743 BP.
XX
AC AAF81712;
XX
DT 01-JUN-2001 (first entry).
XX
DE Human high affinity choline transporter protein encoding cDNA.
XX
KM High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
KN ss.
XX
OS Homo sapiens.
XX
FH Key 1. Location/Qualifiers
FT CDS 1..1743
FT /product= "high affinity choline transporter"
XX
PN MO200116315-A1.
XX
PD 08-MAR-2001.
XX
PF 18-AUG-2000; 2000MWO-JP005545.

```

```

XX 27-AUG-1999; 99UP-00240642.
PR 27-DEC-1999; 99UP-00368991.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PI Haga T, Okuda T;
XX
DR WPI; 2001-226688/23.
XX
DR P-PSDB; AAB74665.
XX
PT New rat and human spinal cord high affinity choline transporters, useful
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
PT treating Alzheimer's disease.
XX
PS Claim 9; Page 71-75; 90pp; Japanese.
XX
CC The present sequence encodes a human (Homo sapiens) high affinity choline
CC transporter protein designated cho-1. The cho-1 protein has nootropic and
CC neuroprotective activities. The cho-1 polynucleotide and protein can be
CC used for the diagnosis of diseases related to the expression of cho-1 by
CC comparing the cho-1 polynucleotide sequence in a sample to that of a
CC control. Drug compositions containing the cho-1 protein or expression
CC promoters or inhibitors of cho-1 are useful for treating disorders
CC characterised by abnormal levels of cho-1, such as Alzheimer's disease
XX
SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,84e-292 Length: 1743
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
DB: Gaps: 0
US-10-724-806-4 (1-580) x AAF81712 (1-1743)
OY 1 MetProPheHisValGlnGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
DB 1 ATGGCTTTCATGATGAGAGAGCTGATGATCATCATCTTCTTACCTTTAATTTGCTG 60
OY 21 ValGlyIleTyrAlaAlaTyrLysThrLysAsnSerGlyAsnProGlnGluArgSerGln 40
DB 61 GTTGAATATGGGCTGCCCTCGAGAGAACCAAAAACAGTGGCAGCGCACAAGCGCACGAA 120
OY 41 AlaIleIleValAlaGlyLysArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
DB 121 GCCATCATGTTGGTGGCCGAGATATTTGTTATTTGTTGGTATTTACATGACAGCT 180
OY 61 ThrTTPValGlyGlyGlyTyrIleAsnGlyThrAlaGlnAlaValTyrGlyProGlyCys 80
DB 181 ACCTGGCTGGAGAGGGATATATCATGTGACAGCTGAAAGCAGTTATATGACAGTTAT 240
OY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
DB 241 GGCTTACCTGGGCTCAGCAGCAATGTGATTTCTTATGTTAGGTGGCTG 300
OY 101 PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln 120
DB 301 TTCTTTCGAAAACCTATGCGTTCAAAAGGGGTATGTGACCATGTTAGACCGTTTCAGCAA 360
OY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlnMetPhe 140
DB 361 ATCTATGAAAACGCTGGGCGGACTCTCTGTTATTTCTTCCGACTGATGGAGAAATGTT 420
OY 141 TTPAlaAlaAlaIlePheSerAlaLeuGlnAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGAGCAATTTCTCTGCTTGGGAGCCACCATCAGCGTATCATCATGTGAT 480
OY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
DB 481 ATGACATTTCTGTATCATCTCTGACATCATTTGACATCTGTATCACACTGTGGAGGG 540

```

```

QY 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db 541 CTCAATTCGTGGCCCTACAGATGATCGTTCAGCTCTTTGGATTCTTTAGGGCGCTGG 600
QY 201 IleserValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThrAlaVal 220
Db 601 ATCAGCGTCCCTTTGGATTGTCACATCTCGACATCGCAGACATGGGGTTCCTGCTGTG 660
QY 221 HisAlaIysTyrGlnSerProTrpLeuGlyThrIleGluSerValGluValTyrThrTrp 240
Db 661 CATGCCAATAACCAAAAGCCGTGGGAGACTGTGACTCATCGAAGTCTACTCTTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrPheGlnArg 260
Db 721 CTGTATGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 261 ValLeuSerSerSerSerSerAlaThrTyrAlaGlnValIleLeuSerPheLeuAlaIlePheGly 280
Db 781 GTTCTCTCTTCTTCTCTCAACCACTATGCTCAAGTGTCTCTCTCTCTCTCTCTCTCTCT 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db 841 TGCCGTGTGATGGCCATCCAGCCATCATTTGGGGCCATTGGAGCATCAACAGACTGG 900
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetIleLeu 320
Db 901 AACCCAGACTGCATATGGGCTTCACATCCAGACTGCAGAAAGAGCAGACATGATTTTA 960
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPheGlyLeuGlyAlaVal 340
Db 961 CCAATTGTTCTGCAGATCTGCCCCGTGTATTTCTTCTGTGTGTGTGTGTGTGTGTGT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1021 TCTGTGCTGTTATCTCATCAACAGATTCTTCCATCTTGTCAAGACTTCCATGTTTGA 1080
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrVal 380
Db 1081 CGGAACATCTACACACTTCTCTTCCAGACAAATGCTTCGGAAGAAATCGTTGGGTT 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db 1141 ATGCGAATCAAGCTTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
QY 401 ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
Db 1201 ACTGTGTAATGGGCTCTGGTACTCAGTCTGACCTTGTATTACATGTTATCTTCCCCAG 1260
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db 1261 CTGCTTGTGTACTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProIle 460
Db 1321 TCTGGCTCTTCTCTAGATAATCACTGAGGGAGGCCAATCTGTATCTTCACTTGCCTT 1380
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnAlaPheProPheLys 480
Db 1381 TTCTACCTCTGCTATTACCTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaIysTyr 500
Db 1441 ACACTTGCATGTATCATCTTCAACCAACATTTGATCTCCATCTTACCAAGATAT 1500
QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaIaArg 520
Db 1501 CTATTGAAAGGGAACCTTGCACCTTAATTAAGATGATGATGATGATGATGATGATGAT 1560
QY 521 HisSerGluGluAsnMetAspLysThrIleLeuValAlaArgAsnGluAsnIleLysLeu 540
Db 1561 CACAGCTGAAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620

```

```

QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db 1621 GAACCTTGACCTTGTGAACCCAGACAGACATGACCTTCACTTCAACATATAA 1680
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
Db 1681 GAGCGCTTCTTGAATGTTGATTCCAGTCCAGAAAGGCTCTGGACATGAAGATTAATTAC 1740

RESULT 11
AAH49207
ID AAH49207 standard; cDNA; 1743 BP.
AC AAH49207;
XX 26-NOV-2001 (first entry)
XX
XX
DE Human CHOT encoding cDNA.
XX
XX CHOT; human; choline transporter; chromosome 2q11-13; nootropic;
XX neuroprotective; gene therapy; antisense therapy; degenerative disease;
XX cognitive disorder; Alzheimer's disease; ss.
OS Homo sapiens.
XX
XX DE1009055-A1.
XX
XX 30-AUG-2001.
XX
XX 28-FEB-2000; 2000DE-01009055.
XX
XX 28-FEB-2000; 2000DE-01009055.
XX
XX 28-FEB-2000; 2000DE-01009055.
XX
XX (BRUE/) BRUESS M.
XX (BOEN/) BOENISCH H.
XX
XX Bruesse M, Boenisch H;
XX
XX WPI; 2001-590709/67.
XX P-PSDB; AAB86837.
XX
XX A new gene encoding human choline transporter, designated hCHOT is
XX located on chromosome 2q11-13 and is useful to treat degenerative
XX disorders such as Alzheimer's disease.
XX
XX Disclosure; Page 11; 12pp; German.
XX
XX This invention describes a novel gene encoding human choline transporter,
XX designated hCHOT which is located on chromosome 2q11-13. The products of
XX the invention have nootropic and neuroprotective activity and can be used
XX for gene or antisense therapy. (1) is used to treat degenerative disease,
XX particularly cognitive disorders such as Alzheimer's disease. Sense and
XX antisense oligonucleotides derived from the gene may be used in
XX diagnostics and other techniques. This sequence encodes the human CHOT
XX protein described in the invention
XX
XX
SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,84e-292 Length: 1743
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
DB: 5 Gaps: 0

US-10-724-806-4 (1-580) x AAH49207 (1-1743)
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuIleuIlePheLeu 20
Db 1 ATGGCTTCCATGTGGAAGGAGCTAGTACATCATCGTCTTACCTTCAATTTTGTCTG 60
QY 21 ValGlyIleTTPAlaAlaIleTyrLysAsnSerGlyAsnProGluGluArgSerGlu 40

```


XX Wu D, Gu Y, Millard WJ, He Y;
 XX WPI: 2003-361535/34.
 DR P-PSDB; AB089879.
 XX
 PT Novel isolated polynucleotide (I) that encodes high affinity choline
 transporter protein, useful for preventing, treating or ameliorating
 PT neurological and cognitive disorders such as Alzheimer's or Parkinson's
 PT disease.

Claim 2; Col 17-21; 20pp; English.

XX The invention relates to an isolated polynucleotide which encodes a high
 CC affinity choline transporter (HACT) protein appearing as AB089879. Also
 CC included are a polynucleotide encoding a fragment consisting of at least
 CC about 50 amino acids of the HACT protein, a vector comprising the
 CC polynucleotide, a composition comprising a vector comprising a
 CC polynucleotide which comprises at least about 12 contiguous nucleic acids
 CC of a polynucleotide appearing as ABX94339 (encoding choline
 CC acetyltransferase), a recombinant host cell which comprises the vector
 CC (used to express the HACT protein or fragment). The polynucleotide is
 CC useful as a probe or primer to detect the presence of HACT polynucleotide
 CC in a sample, such as a biological sample, or for screening for test
 CC agents which bind to the polynucleotide. A pharmaceutical composition
 CC comprising the polynucleotide is useful for preventing, treating or
 CC ameliorating neurological and cognitive disorders e.g. pain, spasticity,
 CC myoclonus, muscle spasm, muscle hyperactivity, epilepsy, stroke, head
 CC trauma, neuronal cell death, multiple sclerosis, spinal chord injury,
 CC dystonia, Alzheimer's disease, myasthenia gravis, multi-infarct
 CC dementia, AIDS dementia, Parkinson's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis (ALS), attention deficit disorder, nicotine
 CC addiction, organic brain syndromes, schizophrenia or memory and cognitive
 CC disorders. HACT is thought to be the rate limiting step in cholinergic
 CC neurotransmitter biosynthesis and regeneration (cholinergic transmissions
 CC are crucial to brain functions such as learning and memory). The present
 CC sequence encodes human HACT

XX Sequence 1743 BP; 411 A; 395 C; 405 G; 532 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-84e-292 Length: 1743
 Score: 2795.00 Matches: 537
 Percent Similarity: 96.64 Conservative: 23
 Best Local Similarity: 92.68 Mismatches: 20
 Query Match: 93.4% Gaps: 0
 DB: 9 Indels: 0

US-10-724-806-4 (1-580) x ABX94338 (1-1743)

QY 1 MetProPheHISValGluGlyLeuValAlaIleIleuPheTyrLeuLeuIlePheLeu 20
 DB 1 ATGGCTTCCATGTCGGAAGAGACTGATACCTATCAGCTTCTTAATTTTGGCTG 60
 QY 21 ValGlyIleTTPAlaAlaATripLysThrLysAsnSerGlyAsnProGluGluArgSerGlu 40
 DB 61 GTTGAATATATGGGCTGCTCGAGAACCAAAACACATGCGCGCAAGAGAGCGCGCAA 120
 QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
 DB 121 GCCATCATAGTTGGTGGCGGAGATATTGGTTTATGGTTGATTTACATGACACACT 180
 QY 61 ThrTPValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
 DB 181 ACCGGGTCGAGAGGAGGTATATCATGCAAGCTGGAAGCAGTTATGTCACAGTTAT 240
 QY 81 GlyLeuAlaTTPAlaHISAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
 DB 241 GGCTTACCTTGGGCTCAGGACCAATTTGATATTTCTTAGCTGATTTTAAAGTGGCTG 300
 QY 101 PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln 120
 DB 301 TTCTTTGGAAAAACGATATGCTTCAAAAGGGGTATGTGACCATGTTTGAACCCGTTTCAGCAA 360

QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
 DB 361 ATCTATGAAAAACGATGGGCGACTCTGTTTATTCCTGCACATGATGGAGAAATGTTTC 420
 QY 141 ThrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
 DB 421 TGGGCTGACGAATTTCTGCTTGGAGCCACATCAGCGTATATCATGATGATGAT 480
 QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
 DB 481 ATGCACATTTCTGTCATCATCTCTGCACCTCATTCGCACTGTGACACACTGGTGGAGG 540
 QY 181 LeuTyrSerValAlaTyrThrAspValValGluLeuPheCysIlePheIleGlyLeuTrp 200
 DB 541 CTCTATTTCTGGCCCTTACACGATGTCCTTACAGCTTTTTCATTTTGTAGGGCTGTGG 600
 QY 201 IleserValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
 DB 601 ATACGCTCCCTTTGCAATGTCATCTCGACATCGCAGACATCGGCTTCACTGCTGTG 660
 QY 221 HisAlaLysTyrGlnSerProTrpLeuGlyThrIleGlnSerValGluValTyrThrTrp 240
 DB 661 CATGCCAAATGCCAAAGCCCGTGGTGGAACTGTGACTCATCTGGAAGTCACTCTTGG 720
 QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrPheGlnArg 260
 DB 721 CTGTATATTTTTCCTGTTTATGTCGTGGTGGAATCCATGCGAAGCATCTTTCAGAGG 780
 QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPhePheGlyAlaPheGly 280
 DB 781 GTTCTCTTCTTCTTCTCAGCACTATGCTCAAGTGTCTTCTTCTGCGAGCTTTCGGG 840
 QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
 DB 841 TGCCGTGGAGGCGCATCCAGCCATCTATGCGGCGCATTTGGAGCCTCCACAGACTGG 900
 QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetLeuLeu 320
 DB 901 AACCAAGATGATATGGGCTTCAGATCCCAAGACTCCAGAAAGGCGCAGATGATGTTTAA 960
 QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyValAlaVal 340
 DB 961 CCAATTTGTTCCAGATATCTTGCCTGTGTATATTTCTTCTTGTGGTGGTGCACATT 1020
 QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
 DB 1021 TCTGCTGCTGTATATGATATCAGAGATTTCTTCCATCTTGTGACGCAAGTTCCATGTTTGA 1080
 QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValITrPVal 380
 DB 1081 CGGAACATCTACCAAGCTTTCCTTCAGACAAATGCTTCGGAACAAAGAAATGTTGGGCTT 1140
 QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
 DB 1141 ATGGCAATCAAGTGTGTTGTGTGAGACATCTGCACAGCCACAGGCTTCTGACGAA 1200
 QY 401 ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
 DB 1201 ACTGATGATGGGCTCTGGTACCTCAGTTCTGACCTTGTTTACATCGTATCTTCCCCAG 1260
 QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
 DB 1261 CTGCTTTGTGTACTTGTGTAAAGGAAACCAACCATATGGGCGCGTGGCAGGTTATGTT 1320
 QY 441 PheGlyLeuPheLeuAspIleThrGlyGlyLysProTyrLeuTyrLeuGlnProLeuIle 460
 DB 1321 TCTGACCTCTTCTGAGATATCTGAGGGGAGGCATATCTGTATCTTCAACCCCTTGATC 1380
 QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
 DB 1381 TTCTACCTGCGCTATTTACCCGATGATATATGATATATATATATGAAATTTCCATTTAAA 1440

[illegible]

RESULT 14
ADV77896
ID ADV77896 standard; cDNA; 1743 BP.

DT 24-FEB-2005 (first entry)

XXX,

Human high affinity choline transporter (hCHT) cDNA #1.

Choline transporter; neuromuscular disorder;
autonomic nervous system disorder; central nervous system disorder;
Parkinson's disease; Huntington's chorea; genetic disorder;
Alzheimer's disease; degeneration; neurological disease; schizophrenia;
psychiatric disease; myasthenia gravis; immune disorder; gene therapy;
cns-gen; nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
muscular-gen; neuroleptic; DNA purification; chromosome 2; gene; ss.

Homo sapiens.

Key	Location/Qualifiers	
CDs	1..1743	
FT	/*tag= a	
FT	/product= "Human CHT protein"	
PN	US2004248838-A1.	
PD	09-DEC-2004.	
PF	01-DEC-2003; 2003US-00724806.	
PR	23-JUL-2001; 2001US-00911077.	
XX	(UYVA-) UNIV VANDERBILT.	
PA	Blakely RD, Apparsundaram S, Ferguson S;	
PI	WPI; 2005-020586/02.	
XX	P-PDB; ADV77897.	
DR	GENBANK; AF276871.	
XX		
PT	New human high affinity choline transporter (CHT) cDNA, useful for	
PT	treating a neuromuscular, autonomic or central nervous system disorder,	
PT	including Parkinson's disease, Huntington's disease, Alzheimer's,	
PT	schizophrenia.	
XX		
PS	Example 1, SEQ ID NO 1; 73pp; English.	
XX		
CC	The present invention provides polynucleotides encoding novel high	
CC	affinity choline transporters (CHTs), methods for their use in screening	
CC	and therapy. The invention is useful for treating neuromuscular	
CC	disorders, autonomic or central nervous system disorders such as	
CC	Parkinson's disease, Huntington's disease, Alzheimer's disease,	
CC	schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is	
CC	also useful in gene therapy. The present sequence is human high affinity	
CC	choline transporter (hCHT) cDNA. The current sequence is that of the	
CC	human CHT cDNA which is located on chromosome 2q12. Note: The sequence	
CC	data for this patent did not form part of the printed specification, but	
CC	was obtained in electronic format directly from the US patent office at	
CC	seqdata.uspto.gov/sequence.html?DocID=US2004024838.	
XX		
SEQ	Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;	
XX		
Alignment Scores:		
Pred. No.:	1,84e-292	Length: 1743
Score:	2795.00	Matches: 537
Percent Similarity:	96.6%	Conservative: 23
Best Local Similarity:	92.6%	Mismatches: 20
Query Match:	93.4%	Indels: 0
DB:	14	Gaps: 0
US-10-724-806-4 (1-580) x ADV77896 (1-1743)		
QY	1 MetProPheHisValGIuGIyLeuValAlaIleIleLeuPheTyxDeuLeuIlePheLeu 20	
DB	1 ATGGCTTCCAGTCGGAAGACTGATACATCGTGTCTTACTCTTCAATTGGCTG 60	
QY	21 ValGIyIleTPRAlAlATrIlyysThrIlysaanserGIyAsnProGIuGIuArGserGIu 40	
DB	61 GTTGGAAATATATGGCTGCTCGGAACCAAAAACAGTGGAGCGAGAGAGAGCGACGCGAA 120	

US-10-724-806-4 (1-580) x ADV77896 (1-1743)

Qy 1 MetProPheHisValGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20

Db 1 ATGCGTTTCCATGTGGAAGACTGATAGCATCATCGTGTCTTACCTTCTAATTTTGGTG 60

21 ValGlyIleTrpAlaIleThrIleValSerAsnSerGluValAsnProGluGluArgSerGlu 40

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

01 G1GGHAIAGGCGCC1GGAGAACCAAAACAG1GGCAGCGCAAGAAAGCGCAAGCGAA 12

QY 41 AlaIleIleValGIyGIyArgAspIleGIyLeuLeuValGIyGIyPheThrMetTrpAla 60
 Db 121 GCCATCATAGTGGTGGCCGAGATATTGGTTATTGGTTGGTGGATTTCACATGACGAT 180
 QY 61 ThrTrpValGIyGIyGIyTrpIleAsnGIyThrAlaGluAlaValTrpGIyCys 80
 Db 181 ACCGTGGGTGGAGAGGATATATCAATGGCACAGCTGAAGCAGTTTATATGACCATGTTAT 240
 QY 81 GIyLeuAlaTrpAlaHisAlaProIleGIyTrpSerLeuSerLeuIleuGIyGIyLeu 100
 Db 241 GGCTTACGTTGGGCTCAGGACCAATGGATATCTTATGATCTGATTTAGGTGGCCG 300
 QY 101 PhePheAlaLysProMetArgSerLysGIyTrpValThrMetLeuAspPropheLysGI 120
 Db 301 TCTTTGGCAAAACCTATAGCGTTCAAGGGGTATGATCACATGTTAACCCTTTTCAGCA 360
 QY 121 IleTrpGIyLysArgMetGIyGIyLeuLeuPheIleProAlaLeuMetGIyLysMetPhe 140
 Db 361 ATCTATGGAAAAAGCATGGGCGGACCTCGTTTATCTGCACTGATGGAGAAATGTTTC 420
 QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGIyAlaThrIleSerValIleIleAspValAsp 160
 Db 421 TGGGCTGCGACGACATTTCTCTGTTGGAGCCACATCAGCGTGATCATCGATGTGAT 480
 QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTrpThrLeuValGIyGIy 180
 Db 481 ATGCACATTTCTGTCTCATCTCTGCACTCATTTGCCACTGTGACACACTGGTGGAGAG 540
 QY 181 LeuTrpSerValAlaTrpThrAspValValGluLeuPheCysIlePheIleGIyLeuTrp 200
 Db 541 CTTCTATTCGTGGCTCACACTGATGTCTGACCTCTTTTGCATTTTGTAGGGCTGTGG 600
 QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGIyPheThrAlaVal 220
 Db 601 ATCAGGCTCCCTTTCATATGTCATCCGACGTGCAGACATCGAGTCTACTGCTGTG 660
 QY 221 HisAlaLysTrpGlnSerProTrpLeuGIyThrIleGluSerValGluValTrpThrTrp 240
 Db 661 CATGCCAATACCAAAAGCGGTGGCTGGGAGCTGTGACTCATCTGAAGCTATCTCTGG 720
 QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGIyGIyIleProTrpGlnAlaTrpPheGlnArg 260
 Db 721 CTTGATAGATTTCTGTGTGTGATGTGGGTGGAAATCCCATGCAACATATCTTTCAGAG 780
 QY 261 ValLeuSerSerSerSerAlaThrTrpAlaGluValLeuSerPheLeuAlaIlePheGIy 280
 Db 781 GTTCTCTCTCTCTCCACGACCATATGCTCAAGTCTGCTCTTCCCTGGACGCTTTCGG 840
 QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGIyAlaIleGIyAlaSerThrAspTrp 300
 Db 841 TGCCTGGTATGGCCATCCAGCCATACATCATTTGGGGCCATTTGGACATTAACAGCTGG 900
 QY 301 AsnGIyThrAlaTrpGIyTrpProAspProLysThrLysGluGluAlaAspMetIleLeu 320
 Db 901 AACCCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAAGGCGACATGATTTTA 960
 QY 321 ProIleValLeuGlnTrpLeuCysProValTrpIleSerPhePheGIyLeuGIyAlaVal 340
 Db 961 CCAATGTCTCGAGATCTCTGCCCTGTGTATATTTCTTTCTTGGTGTGAGATT 1020
 QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerMetPheAla 360
 Db 1021 TCTGCTGCTGTTATATGATCAGACAGATTTCTTCATCTTTGCGACAGATTCAGTTTGA 1080
 QY 361 ArgAsnIleTrpGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrpVal 380
 Db 1081 CGGAATCATCTACACAGCTTCTCCCTTCAGCAAAATGCTTCGCAAAATAATCGTTTGGGCT 1140
 QY 381 MetArgIleThrValLeuValPheGIyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
 Db 1141 ATGCGATCATCAGTGTGTGTGTGGAGCATCTGCACAAAGCCATGGCTTGTCTGACGAAA 1200
 QY 401 ThrValTrpGIyLeuTrpTrpLeuSerSerAspLeuValTrpIleIleIlePheProGln 420

Db 1201 ACTGTATAGGCTCTGTACCTCAGTTCGACTGTATTAATCGTTATCTTCCCCAG 1260
 QY 421 LeuLeuCysValLeuPheIleLysGIyThrAsnThrTrpGIyAlaAlaIleGIyTrpIle 440
 Db 1261 CTGCTTTGTACTCTTTGTGTAAAGGAGCAACACTATAGGGGCGCTGGCAGGTTATGT 1320
 QY 441 PheGIyLeuPheLeuArgIleThrGIyGIyLysProTrpLeuTrpLeuGlnProLeuIle 460
 Db 1321 TCTGGCTCTTCTCGAGATTAACCTGAGGGAGGCCATATCTGATATTACGCCCTTGATC 1380
 QY 461 PheTrpProGIyTrpTrpSerAspLysAsnGIyIleTrpAsnGlnArgPhePropheLys 480
 Db 1381 TTCTACCTGCTATTAACCTGATGATATGATATATATATATACAGAAATTCATTAATA 1440
 QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTrpLeuAlaLysTrp 500
 Db 1441 ACACTTGCCATGTGATACATCATTTCTTAACCAACATTTGATCTCTATCTAGCCAAAT 1500
 QY 501 LeuPheGluSerGIyThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520
 Db 1501 CTATTTGAAAGTGAACCTTGGCACTTAATATATATATATATATATATATATATATAT 1560
 QY 521 HisSerGIyLysAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
 Db 1561 CACAGTGAAGAAACATGTGATTAACAATTTCTGTCAAAAATGAAATATTAATTAAGAT 1620
 QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
 Db 1621 GAACCTTGCACTTGTGAAGCAGACAGACATGACCTCAGCTCAACCTTACCAATTA 1680
 QY 561 GluAlaLeuLeuAspValLysSerSerProGluGIySerGIyThrGluAspLeuGln 580
 Db 1681 GAGGCTTCTCTGATGTATTTCCAGTCCAGAAAGGCTTGGGACTGAAGATATATTAAC 1740

RESULT 15
 ADD50646
 ID ADD50646 standard; DNA; 1813 BP.
 AC ADD50646;
 XX
 XX 15-JAN-2004 (first entry)
 DT
 XX
 DE High-affinity choline transporter (CHT) associated DNA sequence #2.
 KW High-affinity choline transporter; CHT; cholinergic function;
 KW Parkinson's disease; Huntington's disease; Alzheimer's disease;
 KW schizophrenia; dysautonomia; myasthenia gravis; brain;
 KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;
 KW neuroprotective; neuroleptic; ds.
 XX
 OS Unidentified.
 XX
 PN US2003114399-A1.
 XX
 PD 19-JUN-2003.
 XX
 PF 23-JUL-2001; 2001US-00911077.
 XX
 PR 23-JUL-2001; 2001US-00911077.
 XX
 PA (BLAK/) BLAKELY R D.
 PA (APPA/) APPARISUNDARAM S.
 PA (FERG/) FERGUSON S.
 PI Blakely RD, Apparsundaram S, Ferguson S;
 DR MPI; 2003-810914/76.
 XX
 PT Novel isolated polynucleotide encoding human or mouse high affinity
 PT choline transporter polypeptide, useful in gene therapy to increase
 PT cholinergic function in a cell of a patient suffering from Alzheimer's
 PT disease.

XX Disclosure; SEQ ID NO 9; 74pp; English.

XX The present invention relates to the isolation of polynucleotide
XX sequences encoding human and mouse high-affinity choline transporter
XX (hCHT and mCHT respectively), and the proteins they encode. The gene
XX encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
XX encoding hCHT is useful for expressing hCHT recombinantly. The hCHT
XX polynucleotide sequence when delivered to a cell, increases cholinergic
XX function in the cell that is in a patient having Parkinson's disease,
XX Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or
XX myasthenia gravis. The hCHT antibody is useful for controlling
XX transporter CHT proteins to the brain, and for treating the above
XX mentioned diseases. The antibody is also useful for diagnosing the above
XX mentioned disorders and to detect the influence of cholinergic
XX signaling. The present DNA sequence of unknown function is provided in
XX the electronic sequence data but is not mentioned in the printed
XX specification. Note: The sequence data for this patent was obtained in
XX electronic format directly from the USPTO web site at seqdata.uspto.gov.

XX Sequence 1813 BP; 440 A; 406 C; 417 G; 550 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,966-292	Length:	1813
Score:	2795.00	Matches:	537
Percent Similarity:	96.6%	Conservative:	23
Best Local Similarity:	92.6%	Mismatches:	20
Query Match:	93.4%	Indels:	0
DB:	10	Gaps:	0

US-10-724-806-4 (1-580) x ADD50646 (1-1813)

Qy 1 MetProphenHisValGluGlyLeuValAlaIleIleLeuPheTyrlleuLeuIlePheIleu 20
Db 19 ATGGCTTCCAGTGGAGAACTGATAGCTATCATGCTTCTTACCTTCTAATTTTGGCTG 78
Qy 21 ValGlyIleThrAlaAlaThrIlySerThrIlySaenSerGlyAAsnProGluGlnAArgSerGlu 40
Db 79 GTTGAATATGAGGCGCTGGAGAAACAACACATGGGACGCGCAAGAGCGGACGCGAA 138
Qy 41 AlaIleIleValAlGlyIlyAArgAspIleGlyLeuLeuValAlGlyIlyPheThMetThrAla 60
Db 139 GGCATCATAGTTGGTGGCGGAGATATGGTTATTTGTTGGTGGATTTACCTGACAGCT 198
Qy 61 ThrTrpValAlGlyIlyGlyTyrlleAenGlyThrAlaGlnAlaValTyrlleProGlyCys 80
Db 199 ACCTGGGTGGAGGAGGATATCATGGACAGCTGAAGCAGTTTATGTACAGTTAT 258
Qy 81 GlyLeuAlaThrAlaHisAlaProIleGlyTyrlleSerIleuSerIleuIleleuGlyIlyLeu 100
Db 259 GAGCTAGCTTGGGCTCAGGACCAATGTGATTTCTTATGTCTAATTTTAAAGTGGCTG 318
Qy 101 PhePheAlaIlyProMetAArgSerIlyGlyTyrlleValAlThMetLeuAAspProPheIlyGln 120
Db 319 TTCCTTGGAAACCTATGCGTTCAAGGGGTATGGACATGTTAGACCCGTTTACAGCA 378
Qy 121 IleTyrllyIlyAArgMetGlyIlyLeuLeuPheIleProAlaLeuMetGlyIlyIleMetPhe 140
Db 379 ATCTATGAGAAACGATGGGCGGACCTCTGTTTATTTCCGTGACATGATGGAGAAATGTTTC 438
Qy 141 ThrAlaAlaAlaIlePheSerAlaIleuGlyAlaThrIleSerValIleIleIleAspValAsp 160
Db 439 TGGGTGGAGCAATTTCTCTGCTTTGGAGCCACCATTCAGGATATCATGATGTGAT 498
Qy 161 ValAsnIleSerValIleValSerAlaIleuIleAlaIleLeuTyrlleThreValAlGlyIly 180
Db 499 ATGCACATTTCTGTCATCATCTCTGCACATTCGCACTCTGTACACACTGGTGGAGGG 558
Qy 181 LeuTyrlleSerValAlaTyrlleThraspValAlGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db 559 CTCTATTTCTGGGCTTACATGATGTCTGTTGAGCTTTTTCATTTTGTAGGGCTGTGG 618
Qy 201 IleSerValProPheAlaLeuSerHisProAlaValAlThrAspIleGlyPheThrAlaVal 220

Db 619 ATCAGCGCTCCCTTGGATTTGTCATCTCTGACGTGCGAGCAATCGAGTTTACCTGCTG 678
Qy 221 HisAlaIlyTyrlleGlnSerProTrpLeuGlyThrIleGlnSerValAlGlnAlaValTyrlleTrp 240
Db 679 CATGCGAAATACCAAAAGCCCTGGCTGGGAACGTGTAGCTCATCGAAGTCTACCTTGG 738
Qy 241 LeuAspAsnPheLeuLeuLeuMetIleuGlyIlyIleProTrpGlnAlaTyrllePheGlnArg 260
Db 739 CTGTATATTTTCTGTTGTTATGATCTGGGTGGAATCCATGAGCAAGCTTACTTTCAGAGG 798
Qy 261 ValIleuSerSerSerSerAlaThrTyrlleAlaGlnValIleuSerPheLeuAlaIlePheIly 280
Db 799 GTTCTCTTCTTCTCTCCAGCACTATGCTCAAGGTGCTCTTCTGCGAGCTTTCGGG 858
Qy 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db 859 TGCCCTGGATGGCATTCGCCCATCTCTATTGGGCGCATTTGAGCATCAACAGACTGG 918
Qy 301 AsnGlnThrAlaTyrllyTyrlleProAspProIlyThrIlyGlnGlnAlaAspMetIleu 320
Db 919 AACGAGACTGATATGGGCTTCCAGATCCCAAGACTACAGAAAGGCGAGACATGATTTTA 978
Qy 321 ProIleValIleuGlnTyrlleuCysProValTyrlleSerPhePheGlyLeuGlyAlaVal 340
Db 979 CCAATTTTTCGAGATATCTCGCCCTGTGTATTTCTTTCTTGGTCTTGGTGCAGTT 1038
Qy 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1039 TCTGCTGCTGTATATGTCATCAGAGATTTCTTCATCTTGTACAGAACTTCATGTTTCA 1098
Qy 361 ArgAsnIleTyrlleGlnLeuSerPheArgGlnAsnAlaSerAspIlyGlnIleValAlTrpVal 380
Db 1099 CGGAACATCTACCGAGCTTCTCTGACAAATGCTTGGCAAAAGAAATGTTGGGCTT 1158
Qy 381 MetArgIleThrAlaIleuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIly 400
Db 1159 ATGCAATCAACAGTGTGTTGTGTGAGCAATCTGCACAGCCATGCGCTTGTACGAAA 1218
Qy 401 ThrValTyrllyLeuTrpTyrlleuSerSerAspIleuValTyrlleIleIlePheProGln 420
Db 1219 ACTGTGATGGGCTCTGGTACTGATCTGACCTGTTTATCATGTTATCTTCCCCAG 1278
Qy 421 LeuLeuCysValIleuPheIleIlyGlyTyrlleThrAsnThrTyrllyAlaValAlGlyTyrlle 440
Db 1279 CTGCTTGTGTACTTGTGTTGAAGGAAACAACACTATGGGCGCGTGGCAGTTATGTT 1338
Qy 441 PheGlyLeuPheLeuArgIleThrGlyIlyIlyProTyrlleuTyrlleuGlyProleuIle 460
Db 1339 TCGGCTCTTCTCGAATATATCGAAGGAGCCATATCTGTATCTTCAAGCCCTTGATC 1398
Qy 461 PheTyrlleProGlyTyrlleTyrlleSerAspIlySaenGlyIlyIleTyrlleGlnArgPheProPheIly 480
Db 1399 TTCTACCTTGGCTATTCACGTATGATATGATATATTAATTCAGAAATTTCCATTTAA 1458
Qy 481 ThrIleuSerMetValThrSerPhePheThrAsnIleCysValSerTyrlleuAlaIlyTyrl 500
Db 1459 ACACCTGGCAGATGATCATCATTTCTTAAACAACATTTGCACCTCTTATCTGACCAAGAT 1518
Qy 501 LeuPheGlnSerGlyThrLeuProProIlyLeuAspValPheAspAlaValAlaIlyArg 520
Db 1519 CTATTTGAAGTGAACCTTCCACCTTAATTAATTAATTAATTAATTAATTAATTAAT 1578
Qy 521 HisSerGlnGlnAsnMetAspIlyThrIleLeuValAArgAsnGlnAsnIleIlyLeuAsn 540
Db 1579 CACAGTGAAGAAACATGATGATTAAGCAATTTCTGTCAAAATATGAATAATTAATTAAT 1638
Qy 541 GlnLeuAlaProValIlyProArgGlnSerIleuThrIleuSerSerThrPheThrAlaIly 560
Db 1639 GAACCTTGACCTTGGAAACCGACAGAGCATGACCCCTCACCTCAACCTTCAACCAATAAA 1698
Qy 561 GlnAlaLeuLeuAspValAspSerSerProGlnIlySerGlyTyrlleGlnAspAsnLeuGln 580

Tue Jul 11 13:42:51 2006

us-10-724-806-4.p2n.rng

Page 21

DB 1699 GAGCCTTCTGATGTGATTCCAGTCCAGAAAGGCTCTGGAGCTGAGATTAATTACAG 1758

Search completed: July 10, 2006, 15:06:15
Job time : 910 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: July 10, 2006, 14:03:36 ; Search time 7655 Seconds
(without alignments)
6355.300 Million cell updates/sec

Title: US-10-724-806-4

Perfect score: 2993
Sequence: 1 MPFHEGVAILFLIFL.....EALIDVDSPEGGSTEDNLQ 580

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 segs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame_plus.model -DEV=xlh
-O=/abs/ABSWEB/spool/US10724806/runat_10072006_140248_7740/app_query.fasta_1
-DB=EST -QFMT=fastad -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs802h
-USER=US10724806@CCN1_1.6323@runat_10072006_140248_7740 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hic:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gasl:*
12: gb_gasl2:*
13: gb_gasl3:*
14: gb_gasl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2973	99.3	4097	6 AK053063	AK053063 Mus muscu
2	2973	99.3	4306	6 AK034415	AK034415 Mus muscu
3	2410	80.5	1743	14 AY413300	AY413300 Mus muscu
4	2257	75.4	1743	14 AY413298	AY413298 Homo sapi

5	2251	75.2	1743	14 AY413299	AY413299 Pan trogl
6	1457	48.7	4122	6 AK164116	AK164116 Mus muscu
7	1243	41.5	882	5 CK770440	CK770440 958613 MA
8	1065	35.6	707	5 CD350164	CD350164 UI-M-FYO-
9	1004	33.5	828	14 CT403380	CT403380 Sus scrofa
10	988.5	33.0	969	4 DM65738	DM65738 CNB333-A0
11	987	33.0	669	4 BY727598	BY727598 BY727598
12	968	32.3	843	10 DT207246	DT207246 JGI_CAA51
13	964	32.2	781	10 DT197042	DT197042 JGI_CAA55
14	953	31.8	1114	10 DP972151	DP972151 CLJ138-H0
15	890	29.7	1054	10 DM614674	DM614674 CLJ282-B0
16	842.5	28.1	672	14 AG157499	AG157499 Pan trogl
17	828	27.7	834	9 CX840109	CX840109 JGI_CAA9
18	824	27.5	516	7 BE233479	BE233479 139685 MA
19	822.5	27.5	800	1 AL669749	AL669749 AL669749
20	790	26.4	548	9 DA361315	DA361315 DA361315
21	784	26.2	576	9 DN991044	DN991044 TC112868
22	746	24.9	753	10 DV838598	DV838598 LB01121.C
23	719	24.0	1021	10 BM002036	BM002036 BM002036
24	719	23.6	1037	1 AL666817	AL666817 AL666817
25	707.5	23.6	934	8 CV123740	CV123740 OSTF30017
26	704	23.5	934	8 BB626260	BB626260 BB626260
27	703	22.9	650	7 CX033587	CX033587 134598 N
28	684	22.8	874	5 CK449221	CK449221 892772 MA
29	681	22.8	874	5 BM629925	BM629925 170006875
30	680.5	22.7	658	3 BM274870	BM274870 BM274870
31	676.5	22.6	549	3 BI630566	BI630566 RH59836.5
32	660.5	22.1	641	2 DN604180	DN604180 ACAC-aab8
33	659	22.0	682	9 CK184985	CK184985 EST774300
34	659	22.0	842	5 BI629504	BI629504 RH58381.5
35	652.5	21.8	640	2 CD360297	CD360297 AGENCOURT
36	638.5	21.3	941	5 DM569404	DM569404 EST_ssa1
37	631.5	21.1	813	10 BY729567	BY729567 BY729567
38	630	21.0	675	4 CX196909	CX196909 Sa_mxo_52
39	629	21.0	658	8 AW668962	AW668962 111664 MA
40	628	21.0	541	7 CK184984	CK184984 EST774299
41	624	20.8	808	5 CB391304	CB391304 OSTF149A8
42	619	20.7	576	4 BJ122485	BJ122485 BJ122485
43	619	20.7	624	2 BJ125564	BJ125564 BJ125564
44	600.5	20.1	565	2 BM277281	BM277281 BM277281
45	600.5	20.1	583	3	

ALIGNMENTS

RESULT 1
AK053063
LOCUS
DEFINITION
Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched library, clone:D930038E20 product:solute carrier family 5 (choline transporter), member 7, full insert sequence.

ACCESSION
AK053063
VERSION
AK053063.1 GI:26343192
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
1 Carninci, P., Hayashizaki, Y., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Komno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system--384-Format
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
PUBMED
11076861

TITLE
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE
AUTHORS

TITLE
The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
REFERENCE
AUTHORS

TITLE
RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.
Antisense transcription in the mammalian transcriptome
Science 309, 1564-1566 (2005)

JOURNAL
REFERENCE
AUTHORS

TITLE
The FANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).
The Transcriptional landscape of the mammalian genome
Science 309, 1559-1563 (2005)

JOURNAL
REFERENCE
AUTHORS

TITLE
Submitted (16-JUN-2001) Yoshinori Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE
Direct Submission
Submitted (16-JUN-2001) Yoshinori Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES
source

1. 4097
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D930038E20"
/db_xref="taxon:10090"
/clone="D930038E20"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="15 days embryo"
512..2254
/note="unnamed protein product; putative

CDS

solute carrier family 5 (choline transporter), member 7
[MGD|MG1:1927126 GB|NM_022025, evidence: BLASTN, 99%,
match=1743]"
/codon_start=1
/protein_id="BAC35253.1"
/db_xref="gi:26343193"
/translation="MSHREBGLVAILFLYLILFLVGIAMKTKNSGNPERSEALIV
GGRDILGVGFMTATWVGGLYINGDEAVYGGCGGLAWAQAFLGISLILGLGLF
AKPMRSKGYVTMLDPFOIYGRKMGGLFLPALMGEMWMAAIFSAIGATISVILIDV
VNISIVSALIAILYTLVGLYSVAIVDVDFICIFIGLMTISVPALSHPAVDIGFT
AVHAKYOSPMILGTIESVEVYTWLNDPLMLMGGLPMDAYFORVSSSAYAOVLSFL
AAPGLVMAIPALICIGALSTDMNNTAYGPDKTEAEAMILPIYQYCPYVYSF
FGLAVASAAWMSADSSILASMSFARITYDLSRKQASDEIYVWVITLVAGASA
TAMALIKTVVGLYLSLDVYIILFPLGLVLTIGNTYGAAGYIFGLFLRTTGG
EPYLYQLPIFPGYPSDKNGIYNORPEPKTISVTSFTNIVSVYAKYLESGTLP
PKLDFPDAVVARHSEENMDKTLIVRENENIKINELAPVAFPROSLTISFTWKREALDV
DSPEGSETEENLQ"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	4097
Score:	2973.00	Matches:	577
Percent Similarity:	99.7%	Conservative:	1
Best Local Similarity:	99.5%	Mismatches:	2
Query Match:	99.3%	Indels:	0
DB:	6	Gaps:	0

US-10-724-806-4 (1-580) x AK053063 (1-4097)

QY	1	MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu	20
DB	512	ATGCTTTCACGAGAGAAAGACGTGACCTATTCCTCTTCTTCTTCTTCTTCTTCTG	571
QY	21	ValGlyIleTyrPalaIleTyrPalaIleTyrPalaIleTyrPalaIleTyrPalaIle	40
DB	572	GTTGGAATATGAGCGCGATGGAACCAAAACAGCGCAACCCAGAGAGCGAGTGA	631
QY	41	AlaIleIleValGlyGlyAlaArgAspIleGlyLeuLeuValGlyGlyPheThrMet	60
DB	632	GCCATCATAGACG	691
QY	61	ThrTyrPalaIleGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGly	80
DB	692	ACCTGGGTTGGAGGCGCTTACATCATAGGACAGAGAACGATGATGAGCGCGGTGT	751
QY	81	GlyLeuAlaTyrPalaIleAlaIleProIleGlyTyrSerLeuSerLeuIleLeuGly	100
DB	752	GCTCAGCTTGGGCGTCAAGCAACCATGATGATCTCGAGCTATTTAGTGTGTCTG	811
QY	101	PhePheAlaIlePheMetArgSerIleGlyTyrValIleMetLeuAspProPheIle	120
DB	812	TTTTTGGGAACCATATCGTTCACAGGATATGACTATGATGATGATGATGATGATG	871
QY	121	IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLeu	140
DB	872	ATCTATGGAAGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	931
QY	141	TyrPalaIleAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspVal	160
DB	932	TGGGCTGAGCAATTTCTTCGATGATGAGGCGCAACATGAGGATGATGATGATGAT	991
QY	161	ValLeuIleSerValIleValIleSerAlaLeuIleAlaIleLeuTyrThrLeuVal	180
DB	992	GGAACATATGCGTATGCTCTGCACTCATGCAATTTTAAACCTATGAGTGGTGG	1051
QY	181	LeuTyrSerValAlaIleTyrThrAspValAlaGluLeuPheCysIlePheIleGly	200
DB	1052	CTCTACTCTGAGCATATATGATGATGATGATGATGATGATGATGATGATGATGAT	1111
QY	201	IleSerValProPheAlaLeuSerHisProAlaValIleAspIleGlyPheThrAla	220
DB	1112	ATCAGTGCCTTTTGGCCCTGTCATCTGAGACGACGACGATGATGATGATGATGAT	1171

Oy 221 Hsajalalytgcinsertprotpleuglythrilegluservalgluvaltyrthrtp 240
 Db 1172 CATGTAATAATACAGAGTCCCTGGCTGGGAAACATTGAAATGAGTGAAGCTACACTGG 1231
 Oy 241 leuapasnphleuleuleuleuwtleuglygllylperotfplnalatyphreglnatg 260
 Db 1232 CTTGAAATATTTCTGTATTTGAGCTGGGGGGAATCCCATGGCAACCTTATCCAGAG 1291
 Oy 261 ValluserSerSerSerSerlaThrYrAlaGlValIleuserPheuluaalaPhegly 280
 Db 1292 GTTCCTCTTTCATCCTCAAGCCACTTGTCTCAAGTCTGTCTTCTTGGGAGCTTTGGG 1351
 Oy 281 CysLeuValMetAlaIleuProAlaIleCysIleGlYAlaIleGlYAlaSerThrAspTP 300
 Db 1352 TGCCTGTGATGGCTCTACCCGCCATATGATAGAGGCTATTGGAGCTTCCACAGCTGG 1411
 Oy 301 AenglnThAlaTyrltyrProapproltyrThrlysgluluaIaspNctIleu 320
 Db 1412 AACCAAGCTCCCTACCGGATCCAGATCCCAAGACTAAGAGAGACAGACATGATTCCT 1471
 Oy 321 ProIleValIleuglnTyreucysProValTyrlleSerPhePheglyleuglyAlaVal 340
 Db 1472 CCGATGCTTCTGAGATCCTTCCCTGTGTACATCTCTTCTTGGGCTTGGTGTCTGT 1531
 Oy 341 SerAlaAlaValMetSerSerAlaAspSerSerIleIleuserAlaSerSerMetPheAla 360
 Db 1532 TAGCTGCTGTACATGCTCCTACAGCTGACTGCTCATCTGTGGGAGTTCATATTGGCT 1591
 Oy 361 ArgAsnIleTyrlgluleuserPheahglAsnAlaSerAspIlysglulIleValITPVal 380
 Db 1592 CGGAATATCTACCGACTTCTTCCCAAGCAAAAGCATCGACAAAGAAATTTGTGTGGCTC 1651
 Oy 381 MetArgIleThrValIleuValPheglYAlaSerAlaThrAlaMetAlaIleuThrIly 400
 Db 1652 ARGAGATACCTGCTTGTGTGGAGCATCTGCACACACCATGGCTTGTCTGACGAAG 1711
 Oy 401 ThrValTyrlglYleuITPtyrIleuserSerAspIleuValTyrlleIleIlePheProgl 420
 Db 1712 ACTGTGTATGGGCTGTGTAACCTGAGCTGAGCTTGTCTACATCATCTTCCACAG 1771
 Oy 421 leuIleCysValIleuPheIlelysglyThrAsnThTyrlglYAlaValaIleTyrlle 440
 Db 1772 CTGCTCTGTGTACTTCTTCAAGAGAACCAACCATTAAGGGGCACTGTGTATTAAT 1831
 Oy 441 PheglYleuPheIleuArgIleThrIlyglYgluProTyrlleuTyrlleuglnProIle 460
 Db 1832 TTGGACATATCTCGAATTTACTGAGAGAGAGCCATATCTATCTTGAGCCCTTAATC 1891
 Oy 461 PheTyrlProglYtyrTyrlSerAspIlyAsnIlyIleTyrlAsnIlyArgPheProPheIly 480
 Db 1892 TTCTACCTCGTATTACTCTGACAGAAATGATATACATAGAGGTTCCCATTTAA 1951
 Oy 481 ThrIleuSerMetValThrSerPhePheThrAsnIleCysValSerTyrlleuAlaIlyTy 500
 Db 1952 ACTCTCTCATGTGTACCTCATCTTACCAACATTTGTCTTCTTATCTAGCAAGTAT 2011
 Oy 501 leuPheglSerIlyThrIleuProProIlyIleuAspValPheAspAlaValAlaIlyArg 520
 Db 2012 CTAATTTGAAAGTGAACCTTGCTCCCAAAATTTGAATTTGATCTGTGTGGCAAG 2071
 Oy 521 HisSerIlygluIAsnMetAspIlyThrIleuValArgAsnIlyAsnIleIlyIleuAsn 540
 Db 2072 CACAGTGAAGAGAACATGAGCAAGACCATCTCATGTAGCAAAATGAAATATCAATTAAT 2131
 Oy 541 GluIleuAlaProValtyrProargIlyIleuIleuThrIleuSerThrPheThrAsnIly 560
 Db 2132 GAATCTGCACCTGTGAAACCTCGCGAGGCTTACCTCACTTCACTTCACTTCACTTCA 2191
 Oy 561 GluAlaIleuIleuAspValAspSerSerProglYIlySerIlyTyrlgluIAsnIleuIle 580
 Db 2192 GAGGCCCTCTTGTATTTGATTTCACTCCGAGAGGGGCTGTGGAGCTGAAGATTAATCA 2251

AK034415
 LOCUS
 DEFINITION
 AK034415
 4306 bp mRNA linear HTC 02-SEP-2005
 Mus musculus adult male diencephalon cDNA, RIKEN full-length
 enriched library, clone:933018BK24 product:solute carrier family 5
 (cholesterol transporter), member 7, full insert sequence.
 AK034415
 VERSION
 AK034415.1 GI:26329926
 HTC, CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2
 10349636
 3
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 11042159
 4
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Katsuna, T., Tachio, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 11076861
 4
 THE RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 THE FANTOM Consortium, the RIKEN Genome Exploration Research Group
 Phase I and II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6
 RIKEN Genome Exploration Research Group, Genome Science Group
 (Genome Network Core Team) and the FANTOM Consortium.
 Antisense transcription in the Mammalian Transcriptome
 Science 309, 1564-1566 (2005)
 7
 THE FANTOM Consortium, Riken Genome Exploration Research Group and
 Genome Science Group (Genome Network Project Core Group).
 The Transcriptional Landscape of the Mammalian Genome
 Science 309, 1559-1563 (2005)
 8
 (bases 1 to 4306)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of


```

Db      1714 TTGGACTACTCTCGAATTACTGAGGAGGAGCCATATCTATCTTGAGCCCTTAATC 1773
Qy      461 PheTyProGlyTyTyTySerAspLysAsnGlyTlleTyAsnLlnArgPheProPheLys 480
Db      1774 TTCTACCTCGGTATTAATCTTGACAGAAATGATATACATATCAAGAGGTCCCATTTAA 1833
Qy      481 ThrLeuSerMetValThrSerPheThrAsnLleCysValSerTyThrLeuAlaLysTy 500
Db      1834 ACTCTTCATCGTGTACCTTCATCTTTTCAACAATTTGGTTTCTTATCTAGCAAGTAT 1893
Qy      501 LeuPheGluSerGlyThrLeuProPolysLeuAspValPheAspAlaValAlaArg 520
Db      1894 CTATTGAAAGTGAACCTTGCTCCAAATATGATGATGATGATGATGATGATGATGATG 1953
Qy      521 HisSerGluGluAsnMetAspLysThrLleLeuValArgAsnGluAsnLleLysLeuAsn 540
Db      1954 CACAGTGAAGAGAACATGACAGACCATTTAGTCAAGAAATGAAATATCAAAATTAAT 2013
Qy      541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db      2014 GAACCTTGACACTGTGAACCTCGGACAGCCTTAACCTTCAGTTCAACTTTCACCAATAG 2073
Qy      561 GluAlaLeuLeuAspValaAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
Db      2074 GAGGCCCTCCTTGATGTGATTCACAGTCGAGGAGGCTGTGAGACTGAATATACCTTACAA 2133

RESULT 3
AY413300 1743 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM4844 gene, VIRUAL TRANSCRIPT, partial sequence,
DEFINITION AY413300
VERSION AY413300.1 GI:39769262
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 1743)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Clivello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
Adams,M.D. and Cargill,M.
Interfing nonneutral evolution from human-chimp-mouse orthologous
gene titios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1743)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Clivello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
chem based on alignment.
FEATURES
source
1.1743
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1743
/locus_tag="HCM4844"

ORIGIN
Alignment Scores: 2,33e-255 Length: 1743
Score: 2410.00 Matches: 487
Percent Similarity: 84.0% Conservative: 0
Best Local Similarity: 84.0% Mismatches: 93

```

```

Query Match: 80.5% Indels: 0
DB: 14 Gaps: 0
US-10-724-806-4 (1-580) x AY413300 (1-1743)
Qy      1 MetProPheHisValGluGlyLeuValAlaIlelleuPheTyThrLeuLlePheLeu 20
Db      1 ATGTCTTTCAGGTAAAGAGCTGGAGCTATATCTCTTCACTCTTAATATTCG 60
Qy      21 ValGlyLleTrrAlaAlaTrrLysThrLysAsnSerGlyAsnProGluGluArgSerGlu 40
Db      61 GTTGGAATATGGGCTTCATGAAACCAAAACAGGGGCAACCCAGAGACAGACAGTGA 120
Qy      41 AlaIlelleValGlyGlyArgAspLleGlyLeuLeuValGlyGlyPheThrMetTrala 60
Db      121 GCATCATAGTGGGGGGCGGTGACATGGTTGTTGGTGGTGGTTTACATGACAGNN 180
Qy      61 ThrTrrValGlyGlyGlyTyrlleAsnGlyThrAlaGluAlaValTyrolProGlyCys 80
Db      181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
Qy      81 GlyLeuAlaTrrAlaHisAlaProLleGlyTySerLeuSerLeuLleuGlyGlyLeu 100
Db      241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
Qy      101 PhePheAlaLysProMetArgSerLysGlyTyValThrMetLeuAspProPheLysGln 120
Db      301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
Qy      121 IleTyGlyLysArgMetGlyGlyLeuPheIleProAlaLeuMetGlyLleuMetPhe 140
Db      361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420
Qy      141 TrrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIlelleAspValAsp 160
Db      421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy      161 ValAsnLleSerValIleValSerAlaLeuIleAlaIleLeuTyThrLeuValGlyGly 180
Db      481 GTGAACATATCGGTCTATGTCTCTGACATCTTGCACTTTTAATCCATGGGGGG 540
Qy      181 LeuTySerValaLafTyThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrr 200
Db      541 CTCTACTCTGGGCAATATCTGATGTGTCAGCCTATTCGCACTTTTATAGACCTGTGG 600
Qy      201 lIleSerValProPheAlaLeuSerHisProAlaValThrAspLleGlyPheThrAlaVal 220
Db      601 ATCAGTGCCTCTTTGCTCCCTGTCACATCTGACGTACCGACATCGGATTCACAGCTGTG 660
Qy      221 HisAlaLysTyGlnSerProTrrPleuGlyThrIleGlySerValGluValTyThrTrr 240
Db      661 CATGCTAAATACAGAGTCCCTGGCGGAGACCATTTGATGATCACTTGAAGCTACACTGG 720
Qy      241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrrGlnAlaTyPheGlnArg 260
Db      721 CTTGATTAATTTCTGTTATGATGCTGGGTGATGCCATGCGAACCTCACTCCAGAGG 780
Qy      261 ValLeuSerSerSerSerAlaThrTyrrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db      781 GTTCCTCTTCACCCACGACCATTAATGCTCAGATGCTGCTTCCTCGGAGCTTTTGGG 840
Qy      281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrr 300
Db      841 TGCCTGTGATGCTCTTACCCGCAATATGATGAGAGCTATTTGAGTTTCAACAGACTGG 900
Qy      301 AsnGlnThrAlaTyrrGlyTyrrProAspProLysThrLysGluGluAlaAspMetIleLeu 320
Db      901 AACCAAGACTGCCATACGGGTATCCAGATCCCAAGACTAAGAGAGAACAGACATGATTC 960
Qy      321 ProLleValLeuGlnTyrrLeuCysProValTyrrIleSerPhePheGlyLeuGlyAlaVal 340
Db      961 CCGATGTTCTGAGTACCTGCGCTGCTGATCATCTCTTCTTGGGCTTGCTGCTTT 1020

```

Oy	34	SerialaaiAvalMetSerSerLiAaspSerSerLiLeuSerLiAserSerMetPheAla	360
Db	1021	TCAGCTGCTGCATGCTCTCCAGCTGACCTGCCTCCTGCTGGCGAGATTCTATGTTGGCT	1080
Oy	361	ArgAsnIleIYrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrrVal	380
Db	1081	CGGAATATCTACCAAGCTTTCCTTCGACAAATGATCAGCAAGAAATTTGTGGGCTC	1140
Oy	381	MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys	400
Db	1141	ATGAGGACACATGCTGCTGTGTGGAGCATCTGCACAGCCATGGCTTTCCTGACGAG	1200
Oy	401	ThrValIYrGlyLeuTrpTyrLeuSerSerAspLeuValIYrIleIleIlePheProGln	420
Db	1201	ACTGCTGATGGGCTCTGGTACTGAGCTCTGACCTTCTCTACATCATCTTCCACAG	1260
Oy	421	LeuLeuCyValLeuPheIleLysGlyThrAsnThrIYrGlyAlaValAlaGlyTyrIle	440
Db	1261	CTGCTCTGTGATCTCTTCATCAAAAGGACCAACTTATGGGGCAGTTGCTGGTTATAT	1320
Oy	441	PheGlyLeuPheLeuArgIleThrGlyGlyGlyProTyrLeuGlnProLeuIle	460
Db	1321	TTTGACCTATTCCTGAGAAATTCGAGAGAGCCATATCTATCTTGACCCCTTATTC	1380
Oy	461	PheTyrProGlyTyrTyrSerAspLysAsnGlyIleIYrAsnGlnArgPheProPheLys	480
Db	1381	TTCTACCTCGGTTATATCTCTGACAGAAATGATATATCAATCAGAGTTCCCATTTAA	1440
Oy	481	ThrLeuSerMetValIThrSerPhePheThrAsnIleCyValSerTyrLeuAlaLysTyr	500
Db	1441	ACTCTCTCCATGGTTACCTCATTTCTTACCAACATTTGTGTTTCTTATCTGACCAAGAT	1500
Oy	501	LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg	520
Db	1501	CTATTGGAATGGAACTTGCCTCCCAAAATTAAGATGATTTGATGCTGTTGTGGCAAG	1560
Oy	521	HisSerGlnGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn	540
Db	1561	CACAGTGAAAGAAACATGACACAGCCATTTCTAGTCAGAAATATGAATAATCAATTAAT	1620
Oy	541	GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys	560
Db	1621	GAACTTCGACCTCTGAAACCTCGGACAGGCTTAACCTCAGTTCAACTTCAACCAATAG	1680
Oy	561	GluAlaLeuLeuAspValAspSerSerProGlnGlySerGlyThrGluAspAsnLeuGln	580
Db	1681	GAGGCGCTCTTGATGTTGATTCAGTCGAGGGGGTCTGGGACCTGAAGATATCTTACA	1740
RESULT 4			
AV413298	AV413298	1743 bp	DNA linear GSS 12-DEC-2003
LOCUS	Homo sapiens HCM4844 gene, VIRUTAL TRANSCRIPT, partial sequence,		
DEFINITION	genomic survey sequence.		
ACCESSION	AV413298		
KEYWORDS	AV413298.1 GI:39769260		
SOURCE	GSS		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 1743) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tenenbaum,D.M., Ciavello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smnsky,J.J., Adams,M.D. and Cargill,M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 1743)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tenenbaum,D.M., Ciavello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smnsky,J.J., Adams,M.D. and Cargill,M.		

TITLE	Perriera, S., Wang, G., Zheng, X. H., White, T. J., Sninsky, J. J., Adams, M. D. and Cargill, M.
JOURNAL	Direct Submissions Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
SOURCE	1. 1743
gene	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>1743 /locus_tag="HCM4844"
ORIGIN	
Alignment Scores:	
Pred. No.:	2, 13e-238
Score:	2257.00
Percent Similarity:	81.4%
Best Local Similarity:	77.6%
Query Match:	75.4%
DB:	14 Gaps: 0
US-10-724-806-4 (1-580) x AY413298 (1-1743)	
QY	1 MetProPheHisValGluGlyLeuValAlaIleIlePheTyrLeuLeuIlePheLeu 20
DB	1 ATGGCTTTCCATGATGGAGAGACTGATAGCTATCATCGTTCTTACCTTCAATTTTCTG 60
QY	21 ValGlyIleTPrAlaAlaTPrLysThrIysAsnSerGlyAsnProGluGluArgSerGlu 40
DB	61 GTTGCAATATGCGCTCCGTGGAGAACCAAAAACAGTGGCAGCGCAGAAAGCGCAGAA 120
QY	41 AlaIleIleValGlyIlyArgAspIleGlyLeuValGlyIlyPheThrMetThraAla 60
DB	121 GCCATCATGATGGTGGCCGACATATGATTTATGGTTGGATTTACCATGACAGNN 180
QY	61 ThrTrpValGlyGlyIlyTyrIleAsnGlyThraIleGluAlaValIlyGlyProGlyCys 80
DB	181 NNN 240
QY	81 GlyLeuAlaTPrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleGlyGlyLeu 100
DB	241 NNN 300
QY	101 PhePheAlaLysProMetArgSerIlyGlyTyrValThrMetLeuAspProPheIyGln 120
DB	301 NNN 360
QY	121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
DB	361 NNN 420
QY	141 TyrAlaAlaAlaIlePheSerAlaLeuGlyIleAlaTrIleSerValIleIleAspValAsp 160
DB	421 NNN 480
QY	161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyIly 180
DB	481 ATGCACATTTCTGTATCATCTCTGCACATCTTCATCTGTACACACCTGTGGAGGG 540
QY	181 LeuTyrSerValAlaIlyThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200
DB	541 CTGATATTCGTGGCCACACTGATGTCGTACGCTCTTTGCAATTTTGTAGGGCGTGG 600
QY	201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB	601 ATGACGCTCCCTTTGCATTCATCATCTGTAGACCTGAGCGCAGACATCGGGTTCACTGCTG 660
QY	221 HisAlaIlyTyrGlnSerProTPrLeuGlyThrIleGluSerValGluAlaTyrThrTyr 240
DB	661 CAGGCCAAATACCAAAAGCCGTGGCTGGGAACGTGTGACTCATCTGAAGCTTACTCTTGG 720

```

QY 241 LeuAspAsnPhenLeuLeuMetLeuGlyGlyLeuProTrogInAlaTyrPheGlnArg 260
Db 721 CTTGATAGTTTCTTCTGTTGATGCTGGGATGATCCATGGACAGATACCTTCAGAG 780
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaAlaPheGly 280
Db 781 GTTCTCTCTCTCTCTCCACGACCACTTATGCTCAAGTCTGCTCTCTCTGGAGCTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTTP 300
Db 841 TGCCTGGTATGGCTGATCCAGCAATCACTTGGGGCCATTGGACATCAACAGACTGG 900
QY 301 AangInThrAlaTyrGlyTyrProAspProLysThrGlyGlnAlaAspMetIleLeu 320
Db 901 AACCAAGACTGATATGGCTTCCAGATCCAGACATCAAGAGAGGACAGATGATTTA 960
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCAATTGTTCTGCAATATCTGCTCCCTGATATATTTCTTCTTGGTCTGGTGCAGTT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerMetPheAla 360
Db 1021 TCTGCTGCTGTTATGATCATCAGACAGATCTTCCATCTTGCACAAATTCATGTTGCA 1080
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTTPVal 380
Db 1081 CGGAACATCTACAGCTTCTTCTTCAAGCAAAATGCTTGGACAAAGAAATCGTTGGGTT 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db 1141 ATGCGAATCAGAGTGTGTTGTTGAGAGATCTGCACAGCCATGCTTGCAGAGAA 1200
QY 401 ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIlePheProGln 420
Db 1201 ACTGTGATATGGCTCTGAGTACCTCAGTTCGACTTGTATTCGTATCTTCCCCCAG 1260
QY 421 LeuLeuCysValLeuPheIleLysGlyTyrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db 1261 CTGCTTGTGATCTCTTGTATAGGAAACCAACCTTATGGGGCCGTGGCAGGTTATGTT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGlnProTyrLeuTyrLeuGlnProLeuIle 460
Db 1321 TCTGGCTCTTCTCTGAGAAATACAGAGGGGAGCCATATCTGATCTTCAAGCCCTTGATC 1380
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
Db 1381 TTCTACCTGGCTATTAACCTGATGATATGATATATATATATATATATATATATATATAA 1440
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
Db 1441 AACCTTGCATGATGATCATCATCTTAACCAACATTTGATCTTCAAGCAAGTAT 1500
QY 501 LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520
Db 1501 CTAATTTGAAGTGAAGCTTGGCACTTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
QY 521 HisSerGlyGlnAsnMetAspLysThrIleLeuValArgAsnGlnLysIleLysLeuAsn 540
Db 1561 CACAGTGAAGAAACATGATATAGACAAATCTTGTCAAAAATGATAATTAATTAATTAATTA 1620
QY 541 GlnLeuAlaProAlaLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db 1621 GAACCTTGCACTTGTGAAGCCAGACAGAGATACCTTCACTTCACTTCACTTCACTTAA 1680
QY 561 GlnAlaIleLeuLeuAspValAspSerSerProGlyGlySerGlyTyrGlnAspAsnLeuGln 580
Db 1681 GAGGCTCTCTCTGATGTTGATCTCAGTCCAGAGAGGCTGGGAGCTGAAGATTAATTAACAG 1740

```

RESULT 5
 AY413299 1743 bp DNA linear GSS 12-DEC-2003
 LOCUS Pan troglodytes HCM4844 gene, VIRUAL TRANSCRIPT, partial sequence,

```

ACCESSION      Genomic survey sequence.
VERSION        AY413299.1 GI:39769261
KEYWORDS       GSS.
SOURCE         Pan troglodytes (chimpanzee)
ORGANISM       Pan troglodytes
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                Homiidae; Pan.
REFERENCE      1 (bases 1 to 1743)
                Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
                Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
                Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
                Adams,M.D. and Cargill,M.
                Inferring nonneutral evolution from human-chimp-mouse orthologous
                gene clusters
                Science 302 (5652), 1960-1963 (2003)
JOURNAL        2 (bases 1 to 1743)
PUBMED         14671302
REFERENCE      Clark,A.G., Tanenbaum,D.M., Nielson,R., Thomas,P., Kejariwal,A.,
                Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
                Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
                Adams,M.D. and Cargill,M.
                Direct Submission
                Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
                Rockville, MD 20850, USA
COMMENT        This sequence was made by sequencing genomic exons and ordering
                them based on alignment.
FEATURES       Location/Qualifiers
                source
                  1..1743
                  /organism="Pan troglodytes"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9598"
                  <1..>1743
                  /locus_tag="HCM4844"
ORIGIN
Alignment Scores:
Pred. No.:      9,85e-238      Length:      1743
Score:          2251.00      Matches:      450
Percent Similarity: 81.2%      Conservative: 21
Best Local Similarity: 77.6%      Mismatches: 109
Query Match:    75.2%      Indels:      0
DB:             14      Gaps:      0
US-10-724-806-4 (1-580) x AY413299 (1-1743)
QY 1 MetProPheHisValGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 1 ATGGCTTCCATGATGGAAGAGCTGATGATCATGCTGTTCTCACTTCAATTTGGCTG 60
QY 21 ValGlyIleTTPAlaAlaTTPlysThrLysAsnSerGlyAsnProGlyGlnArgSerGln 40
Db 61 GTTGGAAATATGGCTGCTCCGAGAACCAAAACAGTGGACGCCCAAGAGCCACGAA 120
QY 41 AlaIleIleValAlaGlyLysArgAspIleGlyLeuLeuValAlaGlyPheThrMetTTPAla 60
Db 121 GCCATCATATGTTGGTGGCGAGATATTTGGTTTATTTGTTGGTGGATTTACCATGACANN 180
QY 61 ThrTPValGlyGlyGlyTyrIleAsnGlyThrAlaGlnAlaValTyrGlyProGlyCys 80
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
QY 81 GlyLeuAlaTTPAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
Db 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
QY 101 PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln 120
Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlnMetPhe 140

```


[illegible]

OY	501	LeuphegiusertgylrhtleProptrolyslenuaspValpheaspaIalvalAlarG	520
Db	1501	CTATTGTGAAGTGAACCTTGCCACCCTAAATTAATATATTGGTTGTTCGACA	1560
OY	521	HisserglugluasemecAsplysThrlleuValarsngluasenlleylsenu	540
Db	1561	CACAGTAGAAGAAAACATGGAATAGACCAATCTCTGCAAAATGAATAATATTATGAT	1620
OY	541	GlueulaaprovallyspchoyginserleuthrienserserThrPhetHasuly	560
Db	1621	GAACTTGCACTTGTGAAGCCACGCACAGACGATCCTCGCTCACCTTCACCAATAAA	1680
OY	561	GlualaleuaspValaspSerSeProcgluglYserglyThrgluaspAspleGln	580
Db	1681	GAGGCCCTTCCTTGATGTTGATTCCAGTCCAGANNGCTCTGGACTGAAGATATTTCAG	1740
RESULT 6			
ASU64116			
LOCUS			
DEFINITION	AK164116	4122 bp mRNA linear HTC 21-SEP-2005	
ACCESSION	AK164116	Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length	
VERSION	AK164116.1	enriched library, clone:CS30033B06 product:solute carrier family 5	
KEYWORDS	HTC; CAP trapper.	(choline transporter), member 7, full insert sequence.	
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		
TITLE	Sciurogasthi; Muridae; Muridae; Mus.		
JOURNAL	1		
PUBMED	Carninci, P. and Hayashizaki, Y.		
REFERENCE	High-efficiency full-length cDNA cloning		
AUTHORS	Meth. Enzymol. 303, 19-44 (1999)		
TITLE	10349636		
JOURNAL	2		
PUBMED	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
REFERENCE	Itoh, M., Komou, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.		
AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to		
TITLE	prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,		
TITLE	Komou, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,		
JOURNAL	Sunji, N., Ishii, Y., Nakamura, S., Hazema, M., Nabino, T., Harada, A.,		
PUBMED	Yamanoto, R., Matsunoto, H., Sekaguchi, S., Ikegami, T., Kasaiwaigi, K.,		
REFERENCE	Fujiwako, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M.,		
AUTHORS	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,		
TITLE	Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.		
JOURNAL	RIKEN integrated sequence analysis (RISA) system--384-format		
PUBMED	sequencing pipeline with 384 multicapillary sequencer		
REFERENCE	Genome Res. 10 (11), 1757-1771 (2000)		
AUTHORS	11076861		
TITLE	4		
JOURNAL	Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,		
PUBMED	Arakawa, T., Hara, A., Fukunishi, Y., Komou, H., Adachi, J., Fukuda, S.,		
REFERENCE	Altawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,		
AUTHORS	Salto, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Salto, R.,		
TITLE	Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,		
JOURNAL	Fleischmann, W., Gaasterland, T., Gisi, C., King, B., Kochiwa, H.,		
PUBMED	Kushl, P., Lewis, S., Matsuo, Y., Nikaido, I., Peccole, G., Quackenbush, J.,		
REFERENCE	Schriml, L.M., Staab, F., Suzuki, R., Tomita, M., Wagner, L.,		
AUTHORS	Washio, T., Sakai, K., Okido, T., Futuno, M., Aono, H., Baldarelli, R.,		
TITLE	Barish, G., Blake, J., Boftelli, D., Bojunga, A., Carninci, P.,		
JOURNAL	de Bonaldo, M.F., Brownstein, M.J., Ball, C., Fletcher, C.,		
PUBMED	Fujita, M., Gariboldi, M., Gustinchik, S., Hill, D., Hochmann, M.,		
REFERENCE	Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Machinomi, L.,		
AUTHORS	Mashima, J., Mazzarelli, J., Mondaeys, P., Nordone, P., Ring, B.,		
TITLE	Ringwald, M., Rodriguez, I., Sakamoto, N., Saeki, H., Sato, K.,		
JOURNAL	Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H.,		
PUBMED	Toyo-oaka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,		
REFERENCE			

CONSRMT	Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.
JOURNAL	Riken Genome Exploration Research Group Phase II Team and the FANTOM Consortium
PUBMED	Functional annotation of a full-length mouse cDNA collection Nature 409 (6621), 685-690 (2001)
REFERENCE	11217851
AUTHORS	5
CONSRMT	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, T., Bono, H., Kondo, S., Nikaido, I., Ogasato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Baladrelli, R., Hill, D. P., Bult, C. C., Hume, D. A., Quackenbush, J., Schiml, L. M., Kanapin, A., Matsuda, H., Batilov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusch, V., Chochia, C., Corbani, L. E., Cousins, S., Dalla, E., Dargatzis, T. A., Fletcher, C. F., Forrest, A., Fraser, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, C., Wagner, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Akazawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
CONSRMT	FANTOM Consortium
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420 (6915), 563-573 (2002)
PUBMED	12466851
REFERENCE	6
AUTHORS	Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Fith, M. C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V. B., Brenner, S. E., Batalov, S., Forrest, A. R., Zavolan, M., Davis, M. J., Wilming, L. G., Aidinis, V., Allen, J. E., Ambesi-Impombato, A., Apweiler, R., Attalaya, R. N., Bailey, T. L., Bansal, M., Baxter, L., Belsel, K. W., Bersano, T., Bono, H., Chalk, A. M., Chin, K. P., Choudhary, V., Christofideis, A., Clutterbuck, D. R., Crowe, M. L., Dalla, E., Dalrymple, B. P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C. F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T. R., Gojobori, T., Green, R. E., Guscinich, S., Harbers, M., Hayash, Y., Hensch, T. K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeno, K., Iwama, A., Ishikawa, T., Jaki, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelsø, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S. P., Kruger, A., Kummerfeld, S. K., Kurochkin, I. V., Lareau, L. F., Lazarevic, P., Lipovich, L., Liu, J., Liuni, S., McMilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, K., Morita, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakaguchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K. C., Pavan, W. J., Pavoni, G., Pesole, G., Petrovsky, N., Plaza, S., Reed, J. F., Reid, J. F., Ring, B. Z., Rindwald, M., Rost, B., Ruan, Y., Salzberg, S. L., Sandelin, A., Schneider, C., Schombach, C., Sekiguchi, K., Semple, C. A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sincalati, B., Sperling, S., Stupka, E., Sugtara, K., Sultana, R., Takanaka, Y., Taki, K., Tammo, K., Tan, S. L., Tang, S., Taylor, M. S., Tegner, J., Teichmann, S. A., Ueda, H. R., van Nimwegen, E., Verardo, R., Wei, C. L., Yagi, K., Yamashita, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S. M., Teasdale, R. D., Liu, E. T., Brusch, V., Quackenbush, J., Wahlestedt, C., Mattick, J. S., Hume, D. A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M.,
CONSRMT	Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Komno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanuki, A., Okamura-Oho, Y., Suzuki, H., Kawai, T. and Hayashizaki, Y.
JOURNAL	FANTOM Consortium
PUBMED	The transcriptional landscape of the mammalian genome Science 309 (5740), 1559-1563 (2005)
REFERENCE	16141072
AUTHORS	7
CONSRMT	Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakatani, M., Nakamura, M., Nishida, H., Yeo, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Fith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batilov, S., Engstrom, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.
JOURNAL	Riken Genome Exploration Research Group
PUBMED	Antisense transcription in the mammalian transcriptome Science 309 (5740), 1564-1566 (2005)
REFERENCE	16141073
AUTHORS	8 (bases 1 to 4122)
CONSRMT	Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Horii, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Komno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watanuki, A., Muramatsu, M. and Hayashizaki, Y.
JOURNAL	Direct Submition
PUBMED	Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-research@riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
REFERENCE	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
COMMENT	Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers
FEATURES	1..4122
source	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM_DB:C530033E06"
	/db_xref="taxon:10090"
	/clone="C530033E06"
	/tissue_type="spinal cord"
	/clone_id="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="12 days embryo"
	<279..1160
	/note="unnamed protein product; putative
	solute carrier family 5 (cholesterol transporter), member 7
	(MGD MG1:1927126 GB NM_022025, evidence: BLASTN, 99%,
	match=2369)
	start codon is not identified"
	/codon_start=1
	/protein_id="BAE37635.1"
	/db_xref="GI:74211081"
	/translation="ALSCGFFCSCDMDNTAGYDPKKEADMLIPVLYLCVY
	ISFGLGAVSAWMSADSIISASSMFRKNIYQSFQNSADSKETWWRITLVFG
	ASATMALTLKTVYELWYLSDLVYIIIFPLCLVLFKIGTYGAVAGYIGLFLRI
	TGGEFYLVIQPIIFPYGYSDKNGLYNGFPEKPTLSMTVPTNLCVSLYALPESG
	TLPLPLVDVDAVARQSEENMDKTLVRENKILNELAPVKRQSLTSSFTNNKAL
	IDVDSPPSGEDNMQ"
	4100..4105
	/note="putative"
	4122
	polyA_signal
	polyA_site
CDS	

Alignment Scores:

Pred. No.: 3,95e-149 Length: 4122
 Score: 1457.00 Matches: 255
 Percent Similarity: 97.6% Conservative: 1
 Best Local Similarity: 97.3% Mismatches: 7
 Query Match: 48.7% Indels: 0
 DB: 6 Gaps: 0

US-10-724-806-4 (1-580) x AK164116 (1-4122)

```

OY 288 AAlaiecySilegAlAlaleglYalaserThraspTPrasnGlInThraLaTyrglyTy 307
DB 279 GGTCTCAGTGTGGTGGTTCCTTCGCTCCAGAGCTGGAAACAGACTGCGGTAT 338
OY 308 ProaSPProLYSThrLYsgluGluLaaspMetileLeuProileValleuGlnTyLeu 327
DB 339 CCAGATCCCAAGACTRAGAGGAGACAGACATTCCTCCGATCGTTCTGCAGTACCTC 398
OY 328 CysPProValTYrileSerPhePheglyLeuGlyAlaValSerAlaAlaValMetSerSer 347
DB 399 TGCCCTGGTACATCTCTCTTTGGGCTTGGTGGCTGTTTCAGCTGCTGCATGCTCCA 458
OY 348 AlaasPserSerileLeuSerAlaSerSerMetPheAlaArgAsnilleTyrglinLeuSer 367
DB 459 GCTGACGTGTCATCTCTGTCGGAGAGTTCTTAGTTCCTGGAAATTCACAGACTTTC 518
OY 368 PheArgGlnAsnAlaSerAspLYsgluileValITrpValMetArgileThrValleuVal 387
DB 519 TTCAGACAAATGCATCAGACAGAAATGTGTGGGTCATGAGGATCATCTGCTGTGG 578
OY 388 PheGlyAlaSerAlaThraAlaMetAlaLeuLeuThrylThValTYrGlyLeuTYrTYr 407
DB 579 TTCGAGAGATCTGCAACAGCCATGGCTTTCGAGAAACGTGTATGGCTCTGCTAC 638
OY 408 LeuSerSerAspLeuValTYrileleilePheProGlnLeuLeuCysValleuPheile 427
DB 639 CTGAGCTGACCTGCTCTACATCATCTTCCTCCAGACTGCTGCTGCTTCTTCATC 698
OY 428 LysGlyThraSnrThyrglyAlaValaIaGlyTYrilePheGlyLeuPheLeuArgile 447
DB 699 AAAGAACCAACACTTATGGGGGAGTGTGCTGTATATTTTGGACTATTCCTGGAAT 758
OY 448 ThrGlyGlyGluProTYrleuTYrleuGlnProleuilePheTYrProGlyTYrTYrSer 467
DB 759 ACTGAGAGAGAGCCATATCTATCTATCTGAGCCCTTAATCTTACCTGCTGTAATCT 818
OY 468 AspLYsAsnGlyileTYrAsnGlnArgPheProPheLYSThrLeuSerMetValThrsr 487
DB 819 GACAAAGATGTATATACATCAGAGGTTCCCATTTAAACTCTCTCCATGGTTACCTCA 878
OY 488 PhePheThraSnrileCysValserTYrleuAlaIaLYSTyrLeuPheGluSerGlyThrleu 507
DB 879 TTCTTTACCAACATTTGTTCTTATCTTAGCCAGATATCTATTGAAAGTGGAACTTGG 938
OY 508 ProProlYsLeuAspValPheAspAlaValaIaIaArgHisSerGlyLuanMetLeuP 527
DB 939 CCTCAAAATTAAGATGATTTGATGCTGTTCGCAAGGCAAGTGAAGAACAATGAGC 998
OY 528 LysThriileLeuValArgAsnGluAsnileLYsLeuAsnGluLeuAlaProValLYsPro 547
DB 999 AAGACCATTTCTAGTCAAGAAATTAATAATATCAAAATTAATGAACCTTGCAAACTT 1058
OY 548 ArgGlnSerLeuThrLeuSerSerThrPheThraSnrLYsGluAlaLeuLeuAspValaP 567
DB 1059 CGGCGAGAGCCCTAACCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1118
OY 568 SerSerProGluGlySerGlyThrgLuanAsnLeuGln 580
DB 1119 TTCAGTCCGAGGGGTCTGGGACTGAAGATTAACCTTAACA 1157

```

RESULT 7
, CK770440/c

```

LOCUS CK770440 882 bp mRNA linear EST 20-FEB-2004
DEFINITION 958613 MARC 1BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION CK770440
VERSION CK770440.1 GI:42724534
KEYWORDS EST.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 882)
REFERENCE Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
AUTHORS Caeas,E., Wray,J.B., White,J., Cho,J., Fahrenkrug,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Kohrer,G.A.,
Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F.,
Quackenbush,J. and Keel,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
JOURNAL libraries and construction of a gene index for cattle
PUBMED Genome Res. 11 (4), 626-630 (2001)
11282978
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel.: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plates: 95 row: L column: 20
Seq primer: CTAAACGACCTCAGCATATGAGG.
FEATURES
source
1..882
Location/Qualifiers
/mol_type="mRNA"
/organism="Bos taurus"
/db_xref="taxon:9913"
/life_type="pooled"
/lab_host="DH10B"
/clone_id="MARC 1BOV"
/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
ORIGIN
Alignment Scores:
Pred. No.: 1.88e-126 Length: 882
Score: 1243.00 Matches: 241
Percent Similarity: 96.6% Conservative: 13
Best Local Similarity: 91.6% Mismatches: 9
Query Match: 41.5% Indels: 0
DB: 5 Gaps: 0
US-10-724-806-4 (1-580) x CK770440 (1-882)
OY 318 MetIleuProlileValleuGlnTyrleuCysProValTYrileSerPhePheglyleu 337
DB 880 ATGATCTTGGCGATGTCTCAAGTATCTGCCCCGTAATTTCTTACTTTGGCTT 821
OY 338 GlyAlaValSerAlaAlaValaMetSerSerAlaasPserSerileLeuSerAlaSerSer 357
DB 820 GGAGCGCTTTCGCGCTGTATGCTCCTCAGAGATTTCTTCATCTTGTGCAAGACTTGG 761
OY 358 MetPheAlaArgAsnilleTYrGlnLeuSerPheArgGlnAsnAlaSerAspLYsgluile 377
DB 760 ATGTTGCTCGCAACATCTACAGACTTTCATTCAGACAAATGCTTCGACAAAGAGATA 701
OY 378 ValITrpValMetArgileThrValleuValPheGlyAlaSerAlaThraAlaMetAlaLeu 397
DB 700 GTCTGGGTCATGCGCATCAGGATTTTGTGAGCTTCTGCAATGACATGAGCTTGG 641
OY 398 LeuThrylSerThraValTYrGlyLeuTYrTYrLeuSerSerAspLeuValTYrileleile 417
DB 640 CTAAACGAGAGCGGTATAGGGGCTGTGATCCTGAGCTGCTGAGCTGCTGATCATCATC 581

```

```
QY 418 PheProGlnLeuLeuCyValLeuPheIleuGlyThrAsnThrTyrglyAlaValAla 437
Db 580 TTTCCGCGAGTTGCTGCTGCTGCTTTCATCAGAGGATCCACACGATGAGGGCGGGCA 521
QY 438 G1YTYrIlePheG1YLeuPheLeuArg1IleThrG1YGLuProTyLeuTyLeuGln 457
Db 520 GGTTACATCTCGGGCGCTTTCCGCGAGTCACTGGGGGAGGCGCTACCTGAACCTGCAG 461
QY 458 ProLeuIlePheTyTyProG1YTyTyTySerAspIysAsnG1YIleTyrsnGlnAph 477
Db 460 CTTTGATCTTTTATCCCTGCTATTAATGATTAAGTGCGCATATATACCGAGATTC 401
QY 478 ProPheTyThleuSerMetValThrSerPhePheThrAsnIleCyValSerTyLeu 497
Db 400 CATTATTAACCTTGGCTGCTGCTGATCATATTTTATCCACATTTGCACTTCTTATTA 341
QY 498 AlaTyTyTyLeuPheG1YLeuSerG1YThleuProProTyLeuAspValPheAspAlaVal 517
Db 340 GCCAATATACCTGTTGAAAGTGGAACCTTACACCAAAATTTGATGATTCGATGCTGTT 281
QY 518 ValAlaTyHisSerG1YLeuAsnMetAspIysThrIleuValArgAsnGlnAenIle 537
Db 280 GTCCGAAAGGACAGTGAAGAAAACATGACACACATTCCTGTTAGAAATGAAACAT 221
QY 538 LysLeuAsnGlnLeuAlaProValIysProArgG1YSerLeuThrLeuSerSerThrPhe 557
Db 220 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 161
QY 558 ThrAsnTyG1YAlaLeuLeuAspValAspSerSerProG1YLeuG1YThrg1YAsp 577
Db 160 ACCATCATGAGAGGCGCTCTGGAATGATGATGATGATGATGATGATGATGATGAT 101
QY 578 AsnLeuGln 580
Db 100 AATTATACAG 92

RESULT 8
LOCUS CD350164 707 bp mRNA linear EST 09-JUL-2003
DEFINITION UT-M-F10-cel-h-10-0-UI-r1 NIH_BMAP_F10 Mus musculus cDNA clone
IMAGE:685109 5', mRNA sequence.
ACCESSION CD350164
VERSION CD350164.1 GI:31141679
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS NIH-MGC
TITLE Unpublished (1999)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.niowa.edu/distribution/mouseefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seg primer: pYX-5,
Location/Qualifiers
1. 707
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:685109"
/tissue_type="whole brain"
```

```
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_FY0"
/note="Torgan: Brain; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Heman Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 7 21e-107 Length: 707
Score: 1065.00 Matches: 219
Percent Similarity: 96.9% Conservative: 1
Best Local Similarity: 96.5% Mismatches: 2
Query Match: 35.6% Indels: 5
DB: Gaps: 0

US-10-724-806-4 (1-580) x CD350164 (1-707)
QY 1 MetProPheHisValG1YLeuValAlaIleLeuPheTyLeuLeuIlePheLeu 20
Db 27 ATGCTCTTCCAGTAAAGAGAGCTGAGTATATCTCTTACCTCTTATATTTCTG 86
QY 21 ValG1YIleTyPAlaAlaTyTyThrIysAsnSerG1YAsnProG1YLeuArgSerGln 40
Db 87 GTTGAAATATGGGCTGATGAAACCAAAACAGGGGCAACCAAGAGCGCAGTGAA 146
QY 41 AlaIleIleValG1YGLuValArgAspIleG1YLeuValG1YTherMetThrAla 60
Db 147 GCCATATATGTCGGGGCCGCTGACATTTGTTGTTGGTGTGTTTACATGACAGCC 206
QY 61 ThrTyPValG1YGLuTyTyIleAsnG1YThrAlaGlnAlaValTyrg1YProG1YCy 80
Db 207 ACCTGGGTTTGAAGAGGCTACATCATGAGGACAGCAAGAGATGATGGCCAGGTTGT 266
QY 81 G1YLeuAlaTyPAlaHisAlaProIleG1YTyTySerLeuSerLeuIleLeuG1YLeu 100
Db 267 GGTCTGAGCTTGGGCTCAGGACCCCATTTGATTTCTGAGTAAATTTTAAAGTGTCTG 326
QY 101 PhePheAlaIysProMetArgSerIysG1YTyTyValThrMetLeuAspProPheIysGln 120
Db 327 TTTTTCGCAAACTATGCGTTCCAAAGGATATGATGATGATGATGATGATGATGATG 386
QY 121 IleTyG1YIysArgMetG1YGLuLeuPheIleProAlaLeuMetGly-GluMetPhe 140
Db 387 ATCTATGAAAGAGGAGTGGTGGCTGCTTTCATCCCTGCACTGATGGAGAGATGTT 446
QY 140 eTyPAlaAlaIlePheSerAlaLeuG1YAlaThrIle-SerValIleIle-AspVal 159
Db 447 CTGGGCTGACGACATTTTCTGCACTTAAAGGGGCGACCATTCAGCGATCGAGTGTG 506
QY 160 AspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyTyThrLeuValG1Y 179
Db 507 GATGTAACATATCGGCTATTTGCTGCACTCATTTTATTTATACCTAGTGGGT 566
QY 180 G1YLeuTySerValAlaTyThr-AspValAlaGlnLeuPheCyIlePheIleG1YLe 199
Db 567 GGCTCTACTCTGTGATATATCTGATGATGATGATGATGATGATGATGATGATGATGAT 626
QY 199 UTTP-IleSerValProPheAlaLeuSerHisProAlaValThrAspIleG1YPhThra 219
Db 627 GTGGATCATGATGCTCTTTTGGCTCTGCACTCTGCACTGCACTGCACTGCACTGCACT 686
```


second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5' prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: http://www.openbiosystems.com/cdna_library_construction_fa q.phb8 The primary library was transformed and amplified in DH10B (TI phage resistant) bacteria."

ORIGIN

Alignment Scores:

Pred. No.:	3,67e-98	Length:	988
Score:	988.50	Matches:	192
Percent Similarity:	85.2%	Conservative:	26
Best Local Similarity:	75.0%	Mismatches:	35
Query Match:	33.0%	Indels:	4
	10	Gaps:	1

US-10-724-806-4 (1-580) x DW655738 (1-988)

```

QY 1 MetProPhHisValGluGlyLeuValAlaIleLeuPheTyrlleuLeuIlePheLeu 20
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 223 ATGACCATCATACATAGAGGGGCTTGCTGCTATCGCATCTTCTATCATGATCGTGC 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 21 ValGlyIleThrAlaIleThrIleValSerGlyValSerProGlu-----GluArg 38
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 283 GTGGGCGATCGTGGGCGCATGAGAAAACAACACTCCGGGAGGCGGAGGCGACCGC 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 39 SerGluAlaIleIleValGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMet 58
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 343 AGCGAACCACATCATGTCGCTGGGAGGACATGATATTTCTTGTGGATTACCATG 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 59 ThrAlaThrTrpValGlyGlyIleAsnGlyThrAlaGluAlaValIleGlyPro 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 403 ACAGGACCTGGGCTTGAGAGAGATATATATATGATACAGCTAGATGATGATCGCT 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 79 GlyCysGlyLeuAlaThrAlaHisAlaProIleGlyTyrSerIleuSerIleuLeuGly 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 463 GCGTAGCGTTGGCTTGGCTCAAGCTCCCTTGGATATGCACTGCTGTTGTGGG 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 99 GlyLeuPhePheAlaIleValSerMetArgSerIleGlyTyrValThrMetLeuAspProPhe 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 523 GGTCTATTTTTCGCCAAGCCCATCGCTCAAGGGGTTAGCTCAACATGTTGACCGCTTC 582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 119 IysGlnIleIleTyrGlyIleArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyIle 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 583 CAGCAGATCTACGGGAAACGATGGGTGGCTCTCTTCTATCTGACATCATGAGGAG 642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 139 MetPheThrAlaAlaIlePheSerAlaIleGlyAlaThrIleSerValIleIleAsp 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 643 ATCTTCTGTCGGCGCATCTTATCCGCTCGGTGCTAAGTGTGATGTCATCGGAGAC 702
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuIleVal 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 703 ATCAACATTAAGATGTCGGGTATCTACAGCCCTATTCGATATTTTACCTTGCTT 762
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 179 GlyGlyLeuTyrSerValAlaIleTyrThrAspValAlaGlnLeuPheCysIlePheIleGly 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 763 GAGGAGCTTATCTGCTGCTGCTACCGATGTCGTGACAGCTCTTGCATCTTCTTGG 822
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 199 LeuThrIleSerValProPheAlaLeuSerPheProAlaValAlaAspIleGlyPheThr 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 823 CTGTGATACAGGCTCCGCTGCTGCTGACCAACCCGCGGTGCAAGACATCACCATTA 882
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 219 AlaValHisAlaIleTyrIleGlnSerProTyrPheGlyThrIleGlnSerValGluValTyr 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 883 GCGATGAGAGAGGTATACAGTCCCTGGAGAGGAGCATCCAAAGGAGAAC--CG 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 ThrTrpLeuAspAsnPhe--LeuLeuLeuMetLeuGlyIlePro 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB 941 GTGTGATGATATACCTTCTGCTGCTGCTGATGCTGAGNAGATTTCC 986

RESULT 11

BY727598

LOCUS

BY727598

DEFINITION

BY727598

ACCESSION

BY727598

VERSION

BY727598

KEYWORDS

BY727598

SOURCE

BY727598

ORGANISM

BY727598

REFERENCE

BY727598

AUTHORS

BY727598

TITLE

BY727598

JOURNAL

BY727598

PUBMED

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

COMMENT

BY727598

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

FEATURES
 source

Location/Qualifiers
 1..669
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="B73003H24"
 /tissue_type="medulla oblongata"
 /dev_stage="6 days neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 6 days neonate
 medulla oblongata"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5',
 GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. cDNA went through one round of normalization
 to Rot = 20.0 and subtraction to Rot = 459.0. Second
 strand cDNA was prepared with the primer adapter of
 sequence [5' GAGAGAGATCTCGATTAATTAATATCCCCCCCCCC
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a
 modified pBluescript KS(+) after bulk excision from Lambda
 FLC I."

ORIGIN

Alignment Scores:

Pred. No.: 2,93e-98 Length: 669
 Score: 987.00 Matches: 192
 Percent Similarity: 99.0% Conservative: 1
 Best Local Similarity: 98.5% Mismatches: 2
 Query Match: 35.0% Indels: 0
 DB: 4 Gaps: 0

US-10-724-806-4 (1-580) x BY727598 (1-669)

OY 1 MetProPheHisValGlnGlyLeuValAlaIleIleuPheTyrIleuLeuPheLeu 20
 DB 81 ATGTCTTCCAGTGAAGAGACTGTAGCTATATATCTCTTACCTCTTAAATTTCTG 140
 OY 21 ValGlyIleTyrAlaAlaIleTyrIleuValAsnSerGlyAsnProGlnGluArgSerGln 40
 DB 141 GTTGGAATATGGGCTCGCATGAGAAACCAAAACAGGGGCAACCCAGAAAGCGCAGTGA 200
 OY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
 DB 201 GCCATCATATGTCGGGGCCGTGACATGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 260
 OY 61 ThrTTPValGlyGlyTyrTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyGys 80
 DB 261 ACCTGGGTGGAGAGGCTACATCAATGGAGACAGCAAGCAGTATGGGCGCAGGTTGT 320
 OY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleuGlyGlyLeu 100
 DB 321 GGCTACAGCTGGGCTCAGGCAACCATGGATATCTCTAGCTTAATTTAGGTGCTCG 380
 OY 101 PhePheAlaLysProMetArgSerIleGlyTyrValThrMetLeuAspProPheIleGln 120
 DB 381 TTTTTCGCAAACTATGCTTCCAAAGGATATGTACTATGTAGACCATTTCAACAG 440
 OY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140

DB 441 ATCTATGAAAGCCGATGGGCTGCTCTTCATCCCTGACATGATGGAGATGTTTC 500
 OY 141 TTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIlePheValAsp 160
 DB 501 TGGCTGCAAGATTTTCTTCGATTTAGGGGCCACCATCAGGTGATCATTTGATGAT 560
 OY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
 DB 561 GTGAACATATGCTCATTTGCTCGACATTCGACATTCCTTTAAACCTATGATGGTGG 620
 OY 181 LeuTyrSerValAlaIleTyrThrAspValAlaGlnLeuPheCysIle 195
 DB 621 CTTAAGTGTGGCATATATCATGATGTTTCCAGCTATTTGATTT 665

RESULT 12
 DT207246 843 bp mRNA linear EST 16-AUG-2005
 LOCUS DT207246
 DEFINITION UGI CAAS14763.fwd CAAS Pimphales promelas brain 7-8 month adults,
 males and females pooled (H) Pimphales promelas cDNA clone
 CAAS14763 5', mRNA sequence.
 DT207246
 ACCESSION DT207246 GI:72467300
 VERSION DT207246.1 GI:72467300
 KEYWORDS EST.

SOURCE

ORGANISM

Pimphales promelas
 Pimphales promelas
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Pimphales.
 1 (bases 1 to 843)
 Richardson,P., Lucas,S., Rohsard,D., Deter,J.C., Ng,D.C.,
 Brokstein,P. and Lindquist,E.A.
 DOE Joint Genome Institute Pimphales promelas EST project
 Unpublished (2005)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other ESTs: UGI CAAS14763.rev
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 cDNA library Preparation: DOE Joint Genome Institute:
<http://www.jgi.doe.gov>
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Naming Conventions: EST name is generated by the concatenation of
 the UGI Clone Id and the direction of sequencing. The suffix '.fwd'
 indicates a forward sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 Place: CAAS 0153 row: f column: 19
 High quality sequence scop: 754.
 Location/Qualifiers
 1..843

FEATURES
 source

/organism="Pimphales promelas"
 /mol_type="mRNA"
 /db_xref="taxon:90988"
 /clone="CAAS14763"
 /tissue_type="brain"
 /clone_lib="CAAS Pimphales promelas brain 7-8 month
 adults, males and females pooled (H)"
 /note="Vector: pCMVSPORT6; The library was made from dt
 primed cDNA and cloned into Invitrogen vector pCMVSPORT6.
 Poly A RNA were primed with an oligo dt primer (5') ligated
 to a SalI adapter (5'- TCGACCCAGCGCTCG and 5'-
 CGGACGCGTGG) and digested with NotI. cDNA was size
 selected using 1.1% agarose gel electrophoresis (L
 ~0.5-1.2k, M ~1.2-2.5k, H ~2.5k) then ligated into NotI
 and SalI digested pCMVSPORT6 vector. The work was done at
 DOE Joint Genome Institute."

ORIGIN

Alignment Scores:

Pred. No.: 5.37e-96 Length: 843

OY 119 LysGlnIleTyrGlyLysArgMetGlyLeuLeuPheIleProAlaLeuMetGlyGlu 138
Db 444 CAGCAACTGTATGGAAGAAAGATGGCGGTCTGCTATTCATCCCTGCTAATGAGGGGAA 503
OY 139 MetPheTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp 158
Db 504 ATGTTCTGGTCAGGGGCCATCTTATCTGCGCTTGGGGCCACACTANGTGTGATCATTTAC 563
OY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuVal 178
Db 564 ATGCACATCAATATGCTGTGATTAATTCGGCTCTTATTCGACATTTTCTACACGCTGTT 623
OY 179 GlyGlyLeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGly 198
Db 624 GGAGAGCTTACTCAGCGCCCTACACATGATGTTGCCACCTCTTCTGCACTTTTATGAGA 683
OY 199 LeuTPleSerValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThr 218
Db 684 CTGTGGGTGAGCGCTCCCTTTTGCACTCTCAAAACCAAGCTGTGTCTGACATCGGAGTACA 743
OY 219 AlaValIle-----HisAlaLysTyrGlnSerProTyr 228
Db 744 GCAGTGAAGCAGCTCATCTAAT---CAGACTGCATGG 779

RESULT 14
DT972151 1114 bp mRNA linear EST 21-SEP-2005
LOCUS CLJ138-H03 5', mRNA sequence.
DEFINITION CLJ138-H03 5', mRNA sequence.
ACCESSION DT972151
VERSION DT972151.1 GI:76044969
KEYWORDS EST.
SOURCE Gastrosteus aculeatus (three spined stickleback)
ORGANISM Gastrosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphi; Acanthopterygii; Percomorpha; Gastrosteiformes;
Gasterosteidae; Gastrosteus.
1 (bases 1 to 1114)
Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
Schmutz,J. and Myers,R.M.
Expressed sequence tags from Gastrosteus aculeatus
Unpublished (2003)
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@shgc.stanford.edu
Place: 138
High quality sequence stop: 758.
Location/Qualifiers
1. 1114
/organism="Gastrosteus aculeatus"
/mol_type="mRNA"
/strain="Blitrufojordu marine sticklebacks, Iceland"
/db_xref="taxon:69293"
/clone="CLJ138-H03"
/sex="mixed male and female"
/tissue_type="whole larva"
/dev_stage="21 day old larvae collected at Swarnup Stage 30
(J. Embryol. Exp. Morphol 6: 373-383,1958)"
/lab_host="DH10B (T1 phage resistant)"
/note="Vector: Express 1; Total and poly A+ RNA was
isolated from the indicated stickleback tissue, and a cDNA
library was constructed in the Express 1 plasmid vector by
Open Biosystems. First strand cDNA synthesis was primed
with an 54 bp linker primer containing an oligodT sequence
preceded by a synthetic NotI site (first strand primer:
5'-GACGTGTTAGATCGGAGCGCGCC(TT25-3')". Following
second strand synthesis, cDNAs were made blunt at the end

corresponding to the original 5 prime end of mRNA, and
cloned directionally into the NotI and EcoRV sites of
Express 1. Note that the EcoRV site is typically destroyed
in the blunt end cloning, leaving a junction of the form
'xxxAATC' (where is ATC is the second half of the EcoRV
site, and xxx is derived from the cDNA sequence). A map of
the Express 1 vector is available at:
http://www.openbiosystems.com/cdna_library_construction_fa
q.php#8 The primary library was transformed and amplified
in DH10B (T1 phage resistant) bacteria. Clones available
from Open Biosystems:
<http://www.openbiosystems.com/stickleback>

ORIGIN

Alignment Scores:
Pred. No.: 3,82e-94 Length: 1114
Score: 953.00 Matches: 183
Percent Similarity: 80.2% Conservative: 27
Best Local Similarity: 69.8% Mismatches: 50
Query Match: 31.8% Indels: 2
Gaps: 1
DB: 10

US-10-724-806-4 (1-580) x DT972151 (1-1114)

OY 1 MetProPheHisValGlyGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 278 ATGACCATTCACATAGAGGGGCTTGCGCTATCGGACCTTATCTATCTGATCTGTGC 337
OY 21 ValGlyIleTPAlaAlaIleTPleThrIleYsaSerGlyAsnProGlu-----GluArg 38
Db 338 GTGGGCATCTGGCGCGCATGAGAAACAAACACTCCGGGAGGCGGAGCGGACGACCGC 397
OY 39 SerGluAlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMet 58
Db 398 AGCGAAACCATCATGCTGCTGGTGGAGGACATGGATTAATTTGTTGGATTTACATG 457
OY 59 ThrAlaThrTPValGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyPro 78
Db 458 ACAGGACGTGGTGGAGGAGATATTAATGTATACAGCTGATATGTATCTGCCT 517
OY 79 GlyCysGlyLeuAlaTPAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGly 98
Db 518 GCGTACGCTTGGCTTGGGCTCAAGCTCCCTTGGATATGCACTAGCTTCTGTGGGG 577
OY 99 GlyLeuPhePheAlaLysProMetArgSerIleGlyTyrValThrMetLeuAspProPhe 118
Db 578 GGTATATTTTGGCAAGCCCATCGCTACAGAGGTACGTACCATGTGGACCGGTTG 637
OY 119 LysGlnIleTyrGlyLysArgMetGlyLeuLeuPheIleProAlaLeuMetGlyGlu 138
Db 638 CAGCAGATCTACGGGAAACGATGGGTGCGCTCTTCATACCTGCACCTCATGGAGAG 697
OY 139 MetPheTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp 158
Db 698 ATCTTCTGGTCGCGGCCATCTTATCCGCCCTCGGTGCTACTCTGAGTGTATCTGTGAG 757
OY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuVal 178
Db 758 ATNCACATTAAGATGCTCGGTGTTATTCACAGCGCTCATGTCGATATTTTACACCTTGGTT 817
OY 179 GlyGlyLeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGly 198
Db 818 GGNACACTTACTCGGTGCGCTACACCATATGTCGAGAGCTTTTGCAATCTTCTTGGN 877
OY 199 LeuTPleSerValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThr 218
Db 878 CTGTGATCACTAGCTCCGCTTGGTGAACACCCCTGGGTGTGACATACATCACTACT 937
OY 219 AlaValHisAlaLysTyrGlnSerProTyrLeuGlyThrIleGluSerValGluValTyr 238
Db 938 GCAGTGAACAAGTGTACATCGCTTGGAAACCAACCTCCCAAGGACCCCGGG 997
OY 239 ThrTPleuAspAsnPheLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPhe 258

Db 998 GGGGAGACATACCTCCGCTGTGATGCTTGAAGAATCCCTGGGAAGGATTTTC 1057
 QY 259 GlnArg 260
 Db 1058 CAAAAA 1063

RESULT 15
 DM614674
 LOCUS
 DEFINITION CLJ282-809, y1d-s SHGC-CLJ2 Gasterosteus aculeatus cDNA clone
 ACCESSION DM614674
 VERSION DM614674.1 GI:85217654
 KEYWORDS EST.
 SOURCE Gasterosteus aculeatus (three spined stickleback)
 ORGANISM Gasterosteus aculeatus
 Bacteria; Euteleostomi; Euteleostei; Neopterygii; Teleostei; Euteleostei; Neopterygii; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.

REFERENCE 1 (bases 1 to 1054)
 AUTHORS Kingsley, D.M., Peichel, C., Knecht, A., Balabhadra, S., Grimwood, J., Dickson, M., Schmutz, J., and Myers, R.M.
 TITLE Expressed sequence tags from Gasterosteus aculeatus (2004)
 JOURNAL Unpublished (2004)
 COMMENT Contact: Grimwood, Jane
 Stanford Human Genome Center
 Stanford University School of Medicine
 975 S California Ave, Palo Alto, CA 94304, USA
 Tel: 650 320 5917
 Fax: 650 320 5801
 Email: jane@shgc.stanford.edu
 Plate: 282
 High quality sequence stop: 875.
 Location/Qualifiers
 1. 1054
 /organism="Gasterosteus aculeatus"
 /mol_type="mRNA"
 /strain="Bittorfjordur marine sticklebacks, Iceland"
 /db_xref="taxon:69293"
 /clone="CLJ282-809"
 /sex="mixed male and female"
 /tissue_type="whole larva"
 /dev_stage="21 day old larvae collected at Swarnup Stage 30 (J. Embryol. Exp. Morphol 6: 373-383, 1998)"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1ib="SHGC-CLJ2"
 /note="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligodT sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTGATTCAGATCGGAGCGGCCGCTT)25-3'. Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: http://www.openbiosystems.com/cdna_library_construction_fa q.php#8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria."

ORIGIN

Alignment Scores:
 Pred. No.: 3,38e-87 Length: 1054
 Score: 890.00 Matches: 168
 Percent Similarity: 68.8% Conservative: 75
 Best Local Similarity: 47.6% Mismatches: 106

Query Match: 29.7% Indels: 4
 DB: 10 Gaps: 3
 US-10-724-806-4 (1-580) x DM614674 (1-1054)

QY 73 GUAUAAValTYrGlyProGlyCySGlyLeuAlaTPAlaHlaAlaProIleGlyTYrSer 92
 Db 2 GAGGCAACATACACCCGACATGGGCGGCTGGGGCCCTCATGCTGTGCTATGTC 61

QY 93 LeuSerLeuIleLeuGlyLeuPhePheAlaIysProMetArgSerIysGlyTYrVal 112
 Db 62 CTGACCTCTTCTTCTAGTGGGTTTCTTTTTCGACACCTATGAGAGAACATATGTA 121

QY 113 ThrMetLeuAspProPheIysGlnIleTYrGlyIysArgMetGlyIleLeuPheIle 132
 Db 122 ACAATATGAGACCCCTCCACAAAGTATGGGAAGTGTAGAGAGTGTTGATCTT 181

QY 133 ProAlaLeuMetGlyIleMetPheTPAlaAlaIlePheSerAlaLeuGlyAlaThr 152
 Db 182 CCGGCTCTGTGGCTGATGCTCTGTGGGCGACGACACTGTCTGACCTGGTGGGACC 241

QY 153 IleSerValIleIleAspValAspValaIleSerValIleValSerAlaLeuIleAla 172
 Db 242 ATGAGGTATACCTGACCTGACCTGCTACTCTCATATCATCTCTGAGCGGTGGC 301

QY 173 IleLeuTYrThrLeuValGlyIleTYrSerValAlaTYrThrAspValValGlnLeu 192
 Db 302 ATCATCTACACACTGCTGGGGGGGCTCTACTCTGTGGCCATAGGACGTCATCGACTC 361

QY 193 PheCysIlePheIleGlyLeuTrpIleSerValProPheAlaLeuSerHisProAlaVal 212
 Db 362 ATCCATCTTGTGTGCTGCTGGGGTGTGTGCTTCTCTGATGACCAACCTCACTCT 421

QY 213 ThrAspIleGlyPheThrAlaValHisAlaIysTYrGlySerProTrpIleGlyThrIle 232
 Db 422 ATGACATCTCACTGACGCGCTACACAGACGACCTTCCAGGCTCTGGGGGCGACAGTG 481

QY 233 GluSerValGlyValTYrThrTrpLeuAspAsnPheLeuLeuMetLeuGlyIle 252
 Db 482 GAGCTGATGAGGCGCGGCAAGTGCTTCATGATTCATGCTGTGGCTGGGGGGTTTG 541

QY 253 ProTrpGlnAlaTYrPheGlnArgValLeuSerSerSerSerAlaThrTYrAlaGlnVal 272
 Db 542 GCTTACAGGCGCTTATACCAAGAAATCTATCCGCTCTGCTTACACCCAGGCCAAGTT 601

QY 273 LeuSerPheLeuAlaIlePheGlyCysLeuValMetAlaLeuProAlaIleCysIleGly 292
 Db 602 ACCTGCTTGGCTCTCAGGCTTCTGCTGGGTGGGATCCCTTCATCTGTGGGG 661

QY 293 AlaIleGlyAlaSerThrAspTrpAsnGlnThrAlaTYrGlyTYrProAspProLYrThr 312
 Db 662 GCTGTGGCGCCCTTACAGACTGGAATCTCAACAGCTATGGATTACCCACCATATGAA 721

QY 313 LysGlnGlnAlaAspMetIleLeuProIleValLeuGlnIleCysProValTYrIle 332
 Db 722 CGCAACAGGACAGGCTCAATCTCCCATGCGCTCAGATTCCTCACACACCTCACTGTC 781

QY 333 SerPhePheGlyLeuGlyAlaValSerAlaAlaValMetSerSerAlaAspSerIle 352
 Db 782 TCCATCATCGGATGAGTGGAGTGGCGCGCATGTGCTCTCGATGAGTGGCTGCTT 841

QY 353 LeuSerAlaSerSerMetPheAlaArgAsnIleTYrGlnLeuSerPheArgGlnAsnAla 372
 Db 842 CTGTCTCTGTCATCTTCTTCTCAAAATATCTACAGAAACATCCCTCAGGAAGCAGNCA 901

QY 373 SerAspLysGlnIleValTYrValMetArgIleThrValLeuValPheGlyAlaSerAla 392
 Db 902 TCCGAACGTGAGATGATGGGATGATCCGATCCCTGCTGTGATGAGTGGCTGCTGG 961

QY 393 ThrAlaMetAlaLeuLeuThr---LysThrValTYrGlyLeuTrpTYrIleSerSerAsp 411
 Db 962 CACGCCCTACCTTCTGAGACAGCGTTCTGCTTCT---TGGCTCGTGGGCGTGGAC 1015

Oy 412 LeuValTyrIleIleIlePheProGlnLeuCySVal 424
::: |||::: |||
Db 1016 ATGCTACACATCATGTT---CCTCAGCGGTGCTGCGTC 1051

Search completed: July 10, 2006, 16:12:01
Job time : 7697 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocooperation Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: July 10, 2006, 14:15:32 ; Search time 1898 Seconds
(without alignments)
5632.372 Million cell updates/sec

Title: US-10-724-806-4
Perfect score: 2993
Sequence: 1 MPFHEGLVAILFLIFL.....EALDVDSSPESGSTDNLQ 580

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18892170 segs, 6143817638 residues
Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus.model -DEV=xlh
-O=/abss/ABSSWB2.spool/US10724806/runat_10072006_140255_7855/app_query.fasta_1
-DB=Published_Applications_NA_Main -OFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NORM=ext
-HEBFSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=dbse07
-USER=US10724806@CEN_1_1_1675@runat_10072006_140255_7855 -NCPU=6 -ICPU=3
-NO_WMAP -NEE_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA_Main:
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2993	100.0	1743	3	US-09-911-077A-3 Sequence 3, Appl1

2	2993	100.0	1743	3	US-09-911-077A-23	Sequence 23, Appl1
3	2993	100.0	1743	9	US-10-724-806-3	Sequence 3, Appl1
4	2993	100.0	1743	9	US-10-724-806-23	Sequence 23, Appl1
5	2993	100.0	1743	3	US-09-911-077A-5	Sequence 5, Appl1
6	2993	100.0	1743	9	US-10-724-806-5	Sequence 5, Appl1
7	2993	100.0	1743	3	US-09-911-077A-1	Sequence 1, Appl1
8	2993	100.0	1743	9	US-10-724-806-1	Sequence 1, Appl1
9	2993	100.0	1743	3	US-09-911-077A-9	Sequence 9, Appl1
10	2993	100.0	1743	9	US-10-724-806-9	Sequence 9, Appl1
11	2993	100.0	1743	3	US-09-911-077A-7	Sequence 7, Appl1
12	2993	100.0	1743	9	US-10-724-806-7	Sequence 7, Appl1
13	2993	100.0	1743	3	US-11-097-143-41593	Sequence 41593, A
14	2993	100.0	1743	9	US-11-097-143-41594	Sequence 41594, A
15	2993	100.0	1743	3	US-11-005-030-1	Sequence 501, App
16	2993	100.0	1743	9	US-09-911-077A-19	Sequence 19, Appl1
17	2993	100.0	1743	3	US-10-724-806-19	Sequence 19, Appl1
18	2993	100.0	1743	9	US-09-911-077A-14	Sequence 14, Appl1
19	2993	100.0	1743	3	US-10-724-806-14	Sequence 14, Appl1
20	2993	100.0	1743	9	US-10-302-172-932	Sequence 932, App
21	2993	100.0	1743	3	US-10-156-761-1	Sequence 1, Appl1
22	2993	100.0	1743	9	US-10-055-877-116	Sequence 116, App
23	2993	100.0	1743	3	US-10-455-772-911	Sequence 911, App
24	2993	100.0	1743	9	US-10-455-772-917	Sequence 917, App
25	2993	100.0	1743	3	US-10-874-706-56	Sequence 56, App
26	2993	100.0	1743	9	US-10-757-262-93	Sequence 93, Appl1
27	2993	100.0	1743	3	US-10-768-158-45	Sequence 45, Appl1
28	2993	100.0	1743	9	US-09-733-630-1	Sequence 1, Appl1
29	2993	100.0	1743	3	US-09-928-530-3	Sequence 3, Appl1
30	2993	100.0	1743	9	US-10-162-012-28	Sequence 28, Appl1
31	2993	100.0	1743	3	US-10-162-012-26	Sequence 26, Appl1
32	2993	100.0	1743	9	US-10-156-761-5268	Sequence 5268, App
33	2993	100.0	1743	3	US-10-757-262-1	Sequence 1, Appl1
34	2993	100.0	1743	9	US-10-302-172-234	Sequence 234, App
35	2993	100.0	1743	3	US-09-911-077A-14	Sequence 14, Appl1
36	2993	100.0	1743	9	US-10-724-806-14	Sequence 14, Appl1
37	2993	100.0	1743	3	US-10-535-066-53	Sequence 53, Appl1
38	2993	100.0	1743	9	US-10-282-122A-16541	Sequence 16541, A

ALIGNMENTS

RESULT 1
US-09-911-077A-3
; Sequence 3, Application US/09911077A
; Publication No. US20030114339A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELEY, RANDY D.
; APPLICANT: APPASUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:00808
; CURRENT APPLICATION NUMBER: US/09/911, 077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-09-911-077A-3
Alignment Scores: 2.33e-300 Length: 1743
Pred. No.: 1

Score: 2993.00 Matches: 580
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 3 Gaps: 0

US-10-724-806-4 (1-580) x US-09-911-077A-3 (1-1743)

```
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheThrLeuLeuIlePheLeu 20
DB 1 ATGCTTTCACATGTCGAGAGGACTGGTAGCTATTACCTCTTACCTCTTATTTCG 60
QY 21 ValGlyIleTPAlaAlaTrpLysThrLysAsnSerGlyAsnProGluIleArgSerGlu 40
DB 61 GTTGAATATGAGGCTGCATGGAACCAAAAACAGCGCAACCCAGAAAGAGCGCATGAA 120
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
DB 121 GCATCATATGTCGGGCGCGTGAATGGTTGTGGTGGTGGTTTACCATGACAGCC 180
QY 61 ThrTPValGlyGlyGlyTrpIleAsnGlyThrAlaGluAlaValTrpGlyProGlyCys 80
DB 181 ACCTGGGTTGAGAGGAGGCTACATCATGAGCAGCAGAAAGCATGTATGGCCAGGTTGT 240
QY 81 GlyLeuAlaATrPAlaHisAlaProIleGlyTrpSerLeuSerLeuIleLeuGlyGlyLeu 100
DB 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATCTCTGAGTCATATTTAGGTGCTCG 300
QY 101 PhePheAlaLysProMetArgSerLysGlyTrpValIleThrMetLeuAspProPheLysGln 120
DB 301 TTTTGGGAAACCTATGATGCTTCCAGAGATATGATCTATATGACCATTCACAAACG 360
QY 121 IleTrpGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLeuPhe 140
DB 361 ATCTATGAAAGCGCATGGGTGGGCGCTTCATCCCGCATGATGGAGAGAGTGTTC 420
QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGCAGCAATTTCTCTGCATTAAGGGGCCACCATCAGCGTATCATATGATGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTrpThrLeuValGlyGly 180
DB 481 GTGAACATATCGGTATGCTCTGCACTCATTTGCCATTTCTTATACCTTAGGGGTGG 540
QY 181 LeuTrpSerValAlaTrpThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
DB 541 CTCTACTGTCGGCATATCATGTATGTCAGCATATTCGCAATTTTATAGACGTGG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB 601 ATCATGTCCTCTTTGGCCCTGTCACATCTCTGACATCCAGATCGGATTCACAGCTGG 660
QY 221 HisAlaLysTrpGlnSerProTrpLeuGlyThrIleGlnSerValGluValTrpThrTrp 240
DB 661 CATGCTTAATACCAAGTCCCTGGCTGGAGAACATTGATCATGTTGAAGTCTACACCTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTrpPheGlnArg 260
DB 721 CTGTATATTTCTCTGTTATGATGCTGGGTGAATCCCATGGCAAGCCTTATCCAGAGG 780
QY 261 ValLeuSerSerSerSerAlaThrTrpAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 781 GTCTCTCTCTCATCTTCAGCCACCTATGCTCAGAGTCACTGCTTCTGACACCTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
DB 841 TGCCCTGGTATGCTCTACCCGCCCATATGATAGAGCATTTGGAGCTTCCACAGACTGG 900
QY 301 AsnGlnThrAlaTrpGlyTrpProAspProLysTrpLysGluGluAlaAspMetIleLeu 320
DB 901 AACCAAGACTGCTTACGGGTATCCAGATCCCAAGACTAAGAGAGCAATGATTTCTC 960
QY 321 ProIleValLeuGlnTrpLeuCysProValTrpIleSerPhePheGlyLeuGlyAlaVal 340
```

```
DB 961 CGGATCGTTCTGCAGTACCTCGCCCTGTGTACATCTCTTCTTGGGCTGTGCTGT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1021 TGAGCTGCTGTCATGTCCTTCAGCTGACTGCTGCTCATCTCTGTGGCGAGTTCTATGTTGCT 1080
QY 361 ArgAsnIleTrpGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrpVal 380
DB 1081 CGGAATATCTACACAGCTTCTCTTCAGACAAATGATCAGACAAAGAAATGTGTGGGCTC 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
DB 1141 ATGAGATCATCTGTCCTTGTGTGGACATCTGCACAGCCATGCGCTTGTCTGACGAAG 1200
QY 401 ThrValTrpGlyLeuTrpTrpLysSerSerAspLeuValTrpIleIleIlePheProGln 420
DB 1201 ACTGTGATGGGCTTGGTACTGAGCTCTGACCTTGTCTTACATCATATCTTCCACAG 1260
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTrpGlyAlaValAlaGlyTrpIle 440
DB 1261 CTGCTCTGTGTACTCTTCATCAAGAGAACCAACTTATGGGGCAGTTGCTGTATATTT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyLysProLysTrpLeuLeuProLeuIle 460
DB 1321 TTTGACCTATTCCTGAGATTAATGAGAGAGCCATATCTATATCTTACAGCCCTTAATC 1380
QY 461 PheTrpProGlyTrpTrpSerAspLysAsnGlyIleTrpAsnGlnArgPheProPheLys 480
DB 1381 TTCTACCTGCTTATTTACTCTGACAGAAATGATATATCAATCAGAGGTTCCCATTTTAA 1440
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTrpLeuAlaLysTrp 500
DB 1441 ACTCTCTCATAGTACTCATCTTCTTACCAAACTTGTGTTCTTATCTGACCAAGTAT 1500
QY 501 LeuPheGlySerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaLysArg 520
DB 1501 CTATTTGAAAGGGAACCTTGCCTCCAAATTAATGATTAATGATGCTGTGCGCAAGG 1560
QY 521 HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
DB 1561 CACAGTAAAGAGACATGAGACAGACATTTCTATGACAAATGAAATATCAATTAAT 1620
QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
DB 1621 GAACCTGACCTGTGAACCTCGGACAGACCTTAACCTCACTTCAACCTTCAACCAATAG 1680
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
DB 1681 GAGGCCCTCTTGAATGATTCAGTCCGAGAGGGGTCTGGAGACTGAAGATATATTACAA 1740

RESULT 2
US-09-911-077A-23
; Sequence 23, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911.077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1743)
US-09-911-077A-23
```

Alignment Scores:

Pred. No.:	2,336-300	Length:	1743
Score:	2993.00	Matches:	580
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-724-806-4 (1-580) x US-09-911-077A-23 (1-1743)

```

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrlleuLeuIlePheLeu 20
DB 1 AGGCTTTCATGTGGAGAGAGCTGAGTATATATCTCTTACCTCTTATATTTCTG 60
QY 21 ValGlyIleTrrAlaIleTrrPlyThrIleAsnSerGlyAsnProGluGluAsnSerGlu 40
DB 61 GTTGGAAATGGGCTGCATGAAACCAAAACAGCGCAACCCAGAGAGCGGCTGAA 120
QY 41 AlaIleIleValGlyIleArgAspIleGlyLeuLeuValGlyIlePheThrMetTrAla 60
DB 121 GGCATATATGTGGGGGCGGCTGACATTTGTTTGTGGTGGTTTACATGACAGCC 180
QY 61 ThrTrpValGlyGlyIleAsnGlyThrAlaGluAlaValIleIleIleProGlyCys 80
DB 181 ACTGGGTTGAGAGAGCTACATCATGGGACAGAGAGAGAGTATGGGCGAGGTTGT 240
QY 81 GlyLeuAlaIleTrrAlaIleAsnProIleGlyIleSerLeuSerIleuIleuGlyIleu 100
DB 241 GGCTTACGTTGGGCTCATGACCCCATTTGATATTTCTGATCTGATTTTATGAGTGTG 300
QY 101 PhePheAlaIlePheMetArgSerIleGlyIleValIleThrMetLeuAspProPheIleGln 120
DB 301 TTTTTCGCAAACTTATGCGCTTCCAGGATATGATGATGATGATGATGATGATGAT 360
QY 121 IleTyrlleGlyIleArgMetGlyIleLeuLeuPheIleProAlaLeuMetGlyIleMetPhe 140
DB 361 ATCTATGAGAAAGGAGCTAGGCTGCTCTTCATCTTCATCTGACATGATGGAGAGATGTT 420
QY 141 TrrAlaIleAlaIlePheSerAlaLeuGlyIleAlaIleIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGAGCAATTTCTCTGATTTAGGGCCACATGACGCTGATGATGATGATGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleIleuTyrlleuValIleGlyIle 180
DB 481 GTGAAATATTCGCTCTCTCTGCACTATGCACTTCTTTATACCTTACGAGGCTGG 540
QY 181 LeuTyrlleSerValAlaTyrlleAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
DB 541 CTCTACTCTGTGGCATATATCTGATGTTGTCACGCTATTTGCAATTTTATAGGAGCTGTG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValIleAspIleGlyPheThrAlaVal 220
DB 601 ATCAGAGTCCCTTTTCCCTGTCACATCCGTCAGTCAGTCAGTCAGTCAGTCAGTCAGT 660
QY 221 HisAlaIleTyrlleGlnSerProTrpLeuGlyIleThrIleGlySerValGluValIleTyrlle 240
DB 661 CATGCTTAATCCAGAGCTCCCTGCTGGAGACCATTTGAATGATGATGATGATGATGAT 720
QY 241 LeuAspAsnPhelLeuLeuLeuMetLeuGlyIleIleProTrpGlnAlaTyrllePheGlnArg 260
DB 721 CTGTGATTAATTTCTGTTATGATGCTGGGAGAAATCCCATGAGCAACCTTATCTCCAGAG 780
QY 261 ValIleSerSerSerSerAlaTrpTyrlleAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 781 GTTCTCTCTTCATCCCTGACGACCATGCTGACGATCTGCTTCTCTGAGGCTTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleIleGlyAlaSerThrAspTrp 300
DB 841 TGGCTGTGATGCTTACCCGCGCATATGATGATGATGATGATGATGATGATGATGATG 900
QY 301 AsnGlnThrAlaTyrlleGlyIleProAspProGlyIleGlyGluGluAlaIlePheIleLeu 320

```

```

DB 901 AACCAAGCTGCTTACGGGTATCCAGATCCCAAGACTTAAGAGAGAGACATGATTTCTC 960
QY 321 ProIleValLeuGlnTyrlleuCysProValIleIleSerPhePheGlyLeuGlyAlaVal 340
DB 961 CCGATGTTCTGACAGTCTTGGCTGTGTGATCATCTCTTCTTGGGCTGGTGTGCTGTT 1020
QY 341 SerAlaIleValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1021 TCAGTGTGTGATGCTTCACTGATGCTGTCATCTGTCATCTGTCGCGGAGTTCTATGTTGCT 1080
QY 361 ArgAsnIleTyrlleGlnLeuSerPheArgGlnAsnAlaSerAspIleGluIleValIleTrpVal 380
DB 1081 CCGAATATCTTACCAAGCTTCTTCCACAAATATGATGATGATGATGATGATGATGATGAT 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaIleThrAlaMetAlaLeuLeuThrIle 400
DB 1141 ATGAGATCATCTGCTGTTGTTGTTGAGGATCTGACAGAGCCATGCTTGTGCTGACAGAG 1200
QY 401 ThrValTyrlleGlyLeuTrpTyrlleuSerSerAspLeuValTyrlleIleIlePheProGln 420
DB 1201 ACTGTGTATGGGCTGTGATGCTGAGCTGACCTGTCATCATCATCATCTTCCACAG 1260
QY 421 LeuLeuCysValLeuPheIleIleGlyIleThrAsnThrTyrlleAlaValAlaGlyIleIle 440
DB 1261 CTGCTCTGTGATCTTTCATCAAAAGAGAACCACTTATGCGGCGAGTGTGCTGTTATTT 1320
QY 441 PheGlyLeuPheLeuArgIleThrIleGlyIleProTyrlleuTyrlleuGlnProLeuIle 460
DB 1321 TTTGAGCTATCTCTGAAATATCTGAGAGAGAGCCATATCTATCTTGGAGCCCTTAATC 1380
QY 461 PheTyrlleProGlyTyrlleSerAspIleAsnGlyIleTyrlleAsnGlnArgPheProPheIle 480
DB 1381 TTCTACCTGTTTATATCTGACAGAAATGATGATGATGATGATGATGATGATGATGATG 1440
QY 481 ThrLeuSerMetValIleSerPhePheThrAsnIleCysValSerTyrlleuAlaIleTyrlle 500
DB 1441 ACTCTCTCATGATGATCCCATCTTCTTACCAATTTGTTGTTCTTATCTAGCAAGATAT 1500
QY 501 LeuPheGlnSerGlyIleThrIleuProProIleuAspValPheAspAlaValAlaIleArg 520
DB 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTTGATGATGATGATGATGATGATGATGATG 1560
QY 521 HisSerGluGlnAsnMetAspIleThrIleLeuValArgAsnGluAsnIleIleIleIle 540
DB 1561 CACAGTGAAGAGACATGACAAAGACATTTAGTACAGAAATGAAATATCAAAATTAAT 1620
QY 541 GluLeuAlaProValIleProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIle 560
DB 1621 GAACCTGACACTGTGAACCTCGGAGAGCCCTAACCTCAGTTCAACTTCAACCAATTAAG 1680
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlyIleSerGlyIleThrGluAsnLeuGln 580
DB 1681 GAGGCGCTCTTGATGATGATTTCCAGTCCGAGGGGCTGGGACTGAGATTAATTTACAA 1740

```

RESULT 3

```

US-10-724-806-3
; Sequence 3, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APERSONDARUM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT-00808
; CURRENT APPLICATION NUMBER: US/10/724,806
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1743
; TYPE: DNA

```

```
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1743)
US-10-724-806-3

Alignment Scores:
Pred. No.: 2,336-300 Length: 1743
Score: 2993.00 Matches: 580
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-724-806-4 (1-580) x US-10-724-806-3 (1-1743)
QY 1 MetProPheHisValGlyLeuValAlaIleIleLeuPheThrLeuLeuIlePheLeu 20
Db 1 ATGCCTTCCATGTGGAAAGACTGGTAGCTATTATCTCTTACTCTCTTATATTCTG 60
QY 21 ValGlyIleTPAlaIleTPLeuThrIleAsnSerGlyAsnProGluIleArgSerGlu 40
Db 61 GTTGGAAATATGGGCTGCATCGAAACCAAAACACGGCGAACCCAGAAAGCGCAGTGA 120
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
Db 121 GCCATCATATGCGGGGCGCGTGCATGTGGTTGTGGTTGGTTTAACTGACAGCC 180
QY 61 ThrTPValGlyGlyGlyTyrrIleAsnGlyThrIleGluAlaValTyrrGlyProGlyCys 80
Db 181 AACTGGTGTGGAGAGAGCTACATCATGGAGACAGACAGAGCTATAGGCGCAGGTGT 240
QY 81 GlyLeuAlaTPAlaHisAlaProIleGlyTyrrSerLeuSerLeuIleLeuGlyGlyLeu 100
Db 241 GGTCTAGCTGGGCTGCATGACCCCATTTGGATATCTCTGAGCTAATTTTAGTGGTCTG 300
QY 101 PhePheAlaIleProMetArgSerIleGlyTyrrValThrMetLeuAspProPheGln 120
Db 301 TTTTGGGAAACCTATGGGTTCCAAAGGATATGAGCATATTTAGACCATTCATAACG 360
QY 121 IleTyrrGlyIleArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLeuPhe 140
Db 361 ATCTATGGAAGCGCATGGGTGGCTGCTCTTCACTCCGACCTGATGGAGAGATGTTCT 420
QY 141 TPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 421 TGGGCTGCAGCAATTTCTCTGCAATTAGGGGCCACCATCAGCGTATCATGATGTGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrrThrLeuValGlyGly 180
Db 481 GTGAACATATGGTCAATGTCTCTGCACTGCATTCCTTATTAACCTTAGTGGGTGGG 540
QY 181 LeuTyrrSerValAlaIleTyrrThrAspValValGluLeuPheCysIlePheIleGlyLeuTrp 200
Db 541 CTCATCTGTGGCATATACGTATGTGTCCAGCATTTCTGCAATTTTATAGACTGTGG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThrAlaVal 220
Db 601 ATCAGTGTCCCTTTTGGCCCTGTCAATCTGCAGTCCAGCATCGGATTCACACCTGTG 660
QY 221 HisAlaIleTyrrGlnSerProTrpLeuGlyThrIleGluSerValGluValTyrrThrTrp 240
Db 661 CATGCTTAATATCCAGATGCCCTGGGCGGAAACATTAATCACTGGAAGTCTACACCTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaIleTyrrPheGlnArg 260
Db 721 CTGTATATTTTCTGTATATGATGCTGGGTGGAATCCCATGCGCAAGCTCACTCCAGAGG 780
QY 261 ValLeuSerSerSerSerAlaThrTyrrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db 781 GTCTCTCTTATCTCTGACCACTATGCTCAGGATCTGCTCTCTGCGACCTTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
```

```
|||||
Db 841 TGCCCTGGATGGCTCTACCCGCCATATGATAGAGCATATGTGAGCTTCCACACACCTGG 900
QY 301 AsnGlnThrAlaIleTyrrGlyProAspProIleThrIleGluAlaAspMetIleLeu 320
Db 901 AACCGACTGCTTACGGGTATCCAGATCCCAAGACTAAGAGAGAGACAGATGATTTCTC 960
QY 321 ProIleValLeuGlnTyrrLeuCysProValTyrrIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCGATGCTTTCAGTACTCTGCTCCCTGTGATCACTCTCTTGTGGCTGGTCTGTGTT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1021 TCAGCTGCTGCATATCTCTCAGCTGACTGCTTCATCTGCTGGCGAGTCTATGTTTCT 1080
QY 361 ArgAsnIleTyrrGlnLeuSerPheArgGlnAsnAlaSerAspIleGluIleValTPVal 380
Db 1081 CGGAATATCTACAGCTTCTCTTCCAGACAAATGATATGACACAGACATGGCTTGTGAGAG 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400
Db 1141 ATGAGATCAGCTGCTGCTGTGTCGAGACATCTGCACAGCCATGGCTTGTGACAGAG 1200
QY 401 ThrValTyrrGlyLeuTrpTyrrLeuSerSerAspLeuValTyrrIleIleIlePheProGln 420
Db 1201 ACTGTATATGGCTCTGTGTACCTGAGCTGACCTTGTCTATCATCATATCTTCCACAG 1260
QY 421 LeuLeuCysValLeuPheIleIleGlyTyrrAsnThrTyrrGlyAlaValAlaGlyTyrrIle 440
Db 1261 CTGCTCTGTATCTTTCATCAAGAGAACCAACTTATGGGCGAGTGTGCTGTTATATT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProIleLeuTyrrLeuGlnProIleIle 460
Db 1321 TTTGACATTTCTCGAGATTAATCTGAGAGAGAGCCATATCAATCTTCCAGCCTTAATC 1380
QY 461 PheTyrrProGlyTyrrTyrrSerAspIleAsnGlyIleTyrrAsnGlnArgPheProPheIle 480
Db 1381 TTCTACCTGTGTTATCTCTGACAGATGTATATCAATCAAGAGTTCCTCATTTAA 1440
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrrLeuAlaIleTyrr 500
Db 1441 ACTCTCTCAATGTTACTCATTTCTTATCCAAACTTTGTGTTCTTATCTAGCCACAGAT 1500
QY 501 LeuPheGluSerGlyTyrrLeuProProIleIleAspValPheAspAlaValAlaArg 520
Db 1501 CTATTTGAAATGGGAACCTTCCCTCCAAATATGATGATTTGATGCTGTGTCCAGAG 1560
QY 521 HisSerGluGluAsnMetAspIleThrIleLeuValArgAsnGluAsnIleIleIleAsn 540
Db 1561 CACGTGAAGAGAACATGACACAGCATTTCTAGTCAGAAATGAAATATACAAATTAAT 1620
QY 541 GluLeuAlaProValIleProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIle 560
Db 1621 GAACCTTGCACTGTGAACCTCGGAGAGCCTTACCTCACTTCACTTACCAATAAG 1680
QY 561 GluAlaLeuLeuAspValAspSerSerProGluIleSerGlyTyrrGluAspAsnLeuGln 580
Db 1681 GAGGCTCTCTGTGATTTGATTTCCAGTCCGAGGGGTCTGGGACTGAAGATATTTACA 1740

RESULT 4
US-10-724-806-23
; Sequence 23, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:0080US
; CURRENT APPLICATION NUMBER: US/10/724, 806
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911, 077A
; PRIOR FILING DATE: 2001-07-23
```


100

21 CIGAIAMITICIGIAIIGAGCIGGGIGGMAIC

1. TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA

```
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4904
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (224)..(1966)
; US-09-911-077A-5

Alignment Scores:
Pred. No.: 7,87e-296 Length: 4904
Score: 2956.00 Matches: 572
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 98.6% Mismatches: 4
Query Match: 98.8% Indels: 0
DB: Gaps: 0

US-10-724-806-4 (1-580) x US-09-911-077A-5 (1-4904)

QY 1 MetProPheHISValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
DB ATGCTTTTCCATGTGAAGAGACTAGTACGATTATCTGTCTTACTTCTTATATTCTCG 283
QY 21 ValGlyIleTPrAlaAlaTPrIlyThrIlyAsnSerGlyAsnProGluGluArgSerGlu 40
DB GTTGGAATATGTGGCTGCGCATGAGAAAACCAAAACAGCGGTATGCGAAGAACGACGGA 343
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
DB GGCAATCATATGTTGGGGGCGGAGACATTTGGTTGGTTGGTTGGTTTACCATGACAGCC 403
QY 61 ThrTPrValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
DB ACCTGGGTGGAGGAGGATTACATCAACGGGACAGCTGAAAGCAATTATGGGCGAGGTGT 463
QY 81 GlyLeuAlaTPrAlaAlaIleAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
DB GGTTCTAGCTTGGGCTCAGGACCCCATTTGATATCTCTGATGCTGATTTTAAAGTGGCTCG 523
QY 101 PhePheAlaIlyAspPheMetArgSerIlyGlyTyrValThrMetLeuAspProPheIlyGln 120
DB TTTTGTGCAAAACCTATGCGTTCCAAAGGAGATATGTACTATGTATGACCCGTTTCAACAG 583
QY 121 IleTyrGlyIlyArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
DB ACTATATGAAACCGCATGGGTGGGCTGCTGTTCTATCCCTGCACTGATGGGAGAGATGTC 643
QY 141 TPrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB TGGGCTGCGAGCATTTTCTCTGATTAAGGGCTACCATACAGCTATTCATTAATGATGAT 703
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
DB GGGAACATATCGGTCAATGTCTCCGCACTATGGCATTTTATATCCCTCGGGAGAGG 763
QY 181 LeuTyrSerValAlaTyrThrAspValValGluLeuPheCysIlePheIleGlyLeuTPr 200
DB CTCTACTCTGTGGCAATATCTGATGTGTACAGCTATTCGACATTTTATAGGATTTGTGG 823
QY 201 IleSerValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThrAlaVal 220
DB ATCAAGTGTCCATTTGCCCTGTACATCTCTGCACTGATCCGACCATTTGATTCACGTCTGG 883
QY 221 HisAlaIlyTyrGlnSerProIlePheGlyThrIleGlnSerValGluValTyrThrTPr 240
DB CATGCTAAATACCAAGGTCCCTGGCTGGGGAACCATTTGAATCAAGTCAAGTCACTCG 943
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleProIleGlnAlaTyrPheGlnArg 260
```

```
DB 944 CTGATATTTTCTGTTTGTATGTGGGTGATATCCATATGAGCAAGCTTCTCCAGAGG 1003
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB GTTCTCTCTTCAATGCTCAGCGACCTATGCTCAGGAGTGTCTCTTCTGAGCTTTGGG 1063
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
DB TGCCCTGGATGGCTCTACAGCCATTTGCAATGGGGCCATTGGAGCTTCCACACACTGG 1123
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProIlyThrIlyGlnGluAlaAspMetIleLeu 320
DB AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGAGAGAGACATGATTTCTC 1183
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
DB CCGATTTGTTCACAGTACTCTGCGCTGTATGACATTTCTTCTTGGGCTTGTCTGTGT 1243
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB TCTGCTGCTGCATATGCTCGGCTGACATTCATCCATCCATGACGAAGTTCCATGTTTGT 1303
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspIlyGluIleValTPrVal 380
DB CGGAATATCTACAGCTTCTCTTCCAGACAAATGCAATCAGACAGAAATTCGTGGCTC 1363
QY 1364 ATGAGGATCATCTGTGTTGTTGGTTGGACATCTGCACAGCATGCTTGTCTCAGAAAG 1423
QY 401 ThrValIlyTyrGlyLeuTPrIlyLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
DB ACTGTGTATGGGCTGTGGACTGAGCTGTGACTGCTGTGCTCATATCATCTTCCACAG 1483
QY 421 LeuLeuCysValLeuPheIleIlyGlyThrAsnThrTyrGlyAlaValAlaIleGlyTyrIle 440
DB CTGCTCTGTGATCTTCTTATCATCAAGAAACCAACATTTATGGGCGAGTTCGTATAT 1543
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIle 460
DB TTTGGACTTTTCTCTGAGATATCCGAGAGAGGACATATCTATCTTGCAAGCCTTAATC 1603
QY 461 PheTyrProGlyTyrTyrSerAspIlyAsnGlyIleTyrAsnGlnArgPheProPheIly 480
DB TTTCACTCGGTATTTATCTGACCAAGATGGTATATCAATCAAGAGGTTCCCATTTTAA 1663
QY 481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaIlyTyr 500
DB ACTCTCTCCATAGTTTACCTCATTTCTTTACCAACATTTGTGTTTCTTATCTAGCCAA 1723
QY 501 LeuPheGlnSerGlyThrLeuProProIlyLeuAspValPheAspAlaValAlaArg 520
DB CTATTTGAAAGTGAACCTTGCTCCAAATATTAAGATATTTGATGCTGTGTCTCAAGG 1783
QY 521 HisSerGlnGluAsnMetAspIlyThrIleLeuValArgAsnGluAsnIleIlySLeuAsn 540
DB CACAGTGAAGAGAAATGACAAAGCAATTCATGTCAGAAAGAAACATCAAAATTAAT 1843
QY 541 GluLeuAlaProValIlyProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIly 560
DB GAACTTGCACCTGTAAAGCTTGCACAGAGCTTACACCTCAGTTCAACTTTCACCAATAA 1903
QY 561 GluAlaLeuLeuAspValAspSerSerProGlnGlySerGlyThrGluAspAsnLeuGln 580
DB GAGGCTCTCTCTTGAATGATTCCAGTCCAGAGGGATCTGGGACATGAAGATTAACCTTACA 1963

RESULT 6
US-10-724-806-5
; Sequence 5, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
```

APPLICANT: APPASUNDARAM, SUBRAMANIAM
 APPLICANT: FERGUSON, SHAWN
 TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
 FILE REFERENCE: VBLT:008US
 CURRENT APPLICATION NUMBER: US/10/724,806
 PRIOR FILING DATE: 2003-12-01
 PRIOR APPLICATION NUMBER: US/09/911,077A
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 5
 LENGTH: 4904
 TYPE: DNA
 ORGANISM: Rattus norvegicus
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (224)..(1966)
 US-10-724-806-5

Alignment Scores:

Pred. No.:	7,87e-296	Length:	4904
Score:	2956.00	Matches:	572
Percent Similarity:	99.3%	Conservative:	4
Best Local Similarity:	98.6%	Mismatches:	0
Query Match:	98.8%	Indels:	0
Gaps:	9		0

US-10-724-806-4 (1-580) x US-10-724-806-5 (1-4904)

```

OY      1 MetProphHisValGluGlyLeuValAlaIleIleuPheTyrLeuLeuIlePheLeu 20
Db      224 ATGCTTTCATGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 283
OY      21 ValGlyIleTyrAlaIleTyrPheTyrLeuValGlyIlePheTyrLeuValGlyIle 40
Db      284 GTTGGATATGGGCTGATGAGAAACCAAAACAGAGGATATCAGAGAGAGAGAGAGAG 343
OY      41 AlaIleIleValGlyIleGlyArgAspIleGlyLeuLeuValGlyIlePheTyrLeuVal 60
Db      344 GCATCATAGTGGGGGCGGAGACATTTGGTTGTTGGTTGGTTGGTTGGTTGGTTGGTTGG 403
OY      61 ThrTyrValGlyIleGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db      404 ACCTGGGTGGAGAGAGATTACATCAACGGGACAGCTGAGACAGATTATGGGCGAGTGT 463
OY      81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleuGlyIleu 100
Db      464 GGTCTAGCTTGGGCTCAGGACCCCATTTGATTTCTGAGTCTGATTTTAAAGTGGCTTG 523
OY      101 PhePheAlaIleProMetArgSerIleGlyTyrValThrMetLeuAspProPheLeu 120
Db      524 TTTTTCGAAAACCTTATGCGTTCCAGGAGATATGTACATCTATGTATGACCCGTTCAACAG 583
OY      121 IleTyrGlyValArgMetGlyIleLeuLeuPheIleProAlaLeuMetGlyIleuMetPhe 140
Db      584 ATCTATGAGAAAGCGCATGGGTGGCTGCTGCTTCATCCCTGACATGAGGAGAGATGTT 643
OY      141 TTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db      644 TGGGCTGCGAGCAATTTTCTCTGATTAAGGGCTTACATCAGCCATTCATGATGAGAGAT 703
OY      161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleuTyrThrLeuValGlyIle 180
Db      704 GTGAACATATCGGTATGCTTCCGACATCATTTGATTAACCCCTCGTGGAGAGG 763
OY      181 LeuTyrSerValAlaTyrThrAspValValGluLeuPheCysIlePheIleGlyLeuTyr 200
Db      764 CTCTACTCTGTGGCAATATCTGATGTGTACACCTATTTCTGACATTTTATAGATTTGTG 823
OY      201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db      824 ATCAGTGTCCCATTTGCTGCTGATCCATCCGACAGTACCGACATGGATTCACCTGCTGTG 883
  
```

```

OY      221 HisAlaIleTyrGlnSerProTyrLeuGlyThrIleGluSerValGluValTyrThrTyr 240
Db      784 CATGCTAATATACAGAGTCCCTGGCTGGAGACATGATGATGATGATGATGATGATGATGAT 943
OY      241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleIleProTyrGlnAlaTyrPheGlnArg 260
Db      944 CTGTGATATTTTCTGTGTGTGATGCTGGGTGGAAATCCATGGCAACCTTATCTCCAGAGG 1003
OY      261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db      1004 GTCTCTCTTCATTCGTCAGAGACCTATGCTCAGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1063
OY      281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300
Db      1064 TGCCTGTATGCTCTACAGACCATTTGATGGGGCATTTGAGAGCCCTCCACAGACTGG 1123
OY      301 AsnGlnThrAlaTyrGlyTyrProAspProIleThrGlyGluAlaAspMetIleLeu 320
Db      1124 AACCAACTGCAATATGGGTTTCCAGATCCCAAGACCAAGAGAGAGAGAGAGAGAGAGAGAG 1183
OY      321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db      1184 CCGATGTTCTTACAGTACCTCTGCTGTATCATTTCTTCTTGGCTGTGCTGTGCTGTGCTGT 1243
OY      341 SerAlaAlaValMetSerSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db      1244 TCTGCTGCTGATATGCTCTCGGCTGATCATCATCATCATCATCATCATCATCATCATCAT 1303
OY      361 ArgAsnIleTyrGlnLeuSerPheAlaGlnAsnAlaSerAspIleGluIleValTyrVal 380
Db      1304 CGGAATATCTACAGCTTCTCTTCCATGACAAATATGCTACAGACCAAGAAATATGGTGGCT 1363
OY      381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db      1364 ATGAGATCATCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1423
OY      401 ThrValTyrGlyLeuTyrPheTyrLeuSerSerAspLeuValTyrIleIlePheProGln 420
Db      1424 ACTGTATATGGGCTGTGATCTGACAGCTGACCTGTCTATCAATCATCTTCCACAG 1483
OY      421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db      1484 CTGCTCTGTATCTCTTCAATCAAAAGAACCAACATTAATGGGCAATGTGCTGTATAT 1543
OY      441 PheGlyLeuPheLeuArgIleThrGlyIleGlyProTyrLeuTyrLeuGlnProLeuIle 460
Db      1544 TTTGACCTTTCTGAGATTTACCGAGAGAGACCATATCTATCTTGCAGCCCTTATATC 1603
OY      461 PheTyrProGlyTyrTyrSerAspIleTyrIleTyrAsnGlnArgPheProPheLys 480
Db      1604 TTCTACCGGTATTTATACCTGACAAAGATGATATACATCAGAGGTTCCCATTTAA 1663
OY      481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaIleTyr 500
Db      1664 ACTCTCTCATAGTATACCTCATTTCTTACCAATTTGGTTCCATCTCAGCAAGAT 1723
OY      501 LeuPheGluSerGlyThrIleuProIleuAspValPheAspAlaValAlaIleArg 520
Db      1724 CTATTTGAAAGTGAACCTTGCTCCAAATATGATATTTGATGCTGTGCTCAAGG 1783
OY      521 HisSerGluGluAsnMetAspIleThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
Db      1784 CACAGTGAAGAGAACTGAGCAAGAACCATTTAGTCAAGAAATGAAATCAATCAATTAAT 1843
OY      541 GluLeuAlaProValIleProArgIleuSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db      1844 GAACTTGACCTGTAAAGCTCGACAGACCTTAACCTTCACTTCACTTCACTTCACTTCA 1903
OY      561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
Db      1904 GAGGCTCTCTGTATGTTATTTCCAGTCCAGAGGAGATCTGGAGCTGAAGATATCTTACAA 1963
  
```

RESULT 7

```
US-09-911-077A-1
; Sequence 1, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPASUNDARAM, SUBRAMANIAM
; APPLICANT: PERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:00808
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
; US-09-911-077A-1

Alignment Scores:
Pred. No.: 9,496-280 Length: 1743
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
DB: Gaps: 0

US-10-724-806-4 (1-580) x US-09-911-077A-1 (1-1743)

QY 1 MetProPheHisValIGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
DB 1 ATGGCTTCCAGTGGAGGAGCGATGATGCTATCATCTGCTCTACCTTCTAATTTTCTG 60
QY 21 ValGIYIleTPrAlAlaIleTPrLysThrLysAsnSerGIYAsnProGluGluArgSerGI 40
DB 61 GTTGGAAATATGGGCTGGCTGGAGAACAAAGGCGAGCGCAAGGCAAGGCGAGCGAA 120
QY 41 AlaIleIleValIGLYGlyYArgAspIleGIYLeuLeuValIGLYPheThrMetThrala 60
DB 121 GCCATCATATGTTGGTGGCCGAGATATGTTATTGTTGGTGGATTTACATGACAGCT 180
QY 61 ThrTPrValIGLYGlyYTrIleAsnGIYThrAlaGluAlaValTYrgIYProGIYCys 80
DB 181 ACCTGGTCGGAGGAGGATATCAATGGCACAGCTGAAGCAGTTATGATACAGGTTAT 240
QY 81 GlyLeuAlaTPrAlaHisAlaProIleGIYTrSerLeuSerLeuIleLeuGIYGIYLeu 100
DB 241 GGCTTAGCTTGGGCTCAGGACCAATTTGSAATTTCTTAGTCTGATTTTAAAGTGGCTG 300
QY 101 PhePheAlaYsPProMetArGSerIYsGIYTYrValThrMetLeuAspProPheYsGI 120
DB 301 TTCTTTGCAAAACATATGCTTCAAAAGGGGTATGTGACATGTTAGACCGTTTACGAA 360
QY 121 IleTYrgIYLYsArgMetGIYGIYLeuLeuPheIleProAlaLeuMetGIYGIYMetPhe 140
DB 361 ATCTATGCAAAACGATGGCGGAGCTCCCTGTTATTTCCGCACTGATGGGAGAAATGTT 420
QY 141 TrpAlaIleAlaIlePheSerAlaLeuGIYAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGGAGCAATTTCTCTGCTTTGGAGCGCACCATGAGCGTGAATCAATGTGGAT 480
QY 161 ValAsnIleSerValIleValIleSerAlaLeuIleAlaIleLeuTYrThrLeuValIGLY 180
DB 481 ATGCACATTTCTGTCATCATCTCTGCACATCATTTGCCACTCTGTACACACTGGGGAGG 540
QY 181 LeuTYrSerValAlaTYrThrAspValValGluLeuPheCysIlePheIleGIYLeuTPr 200
DB 541 CTCTATTTCTGTGGCCTTACACTGATGTCTGTGAGCTTTTGTGCAATTTTGTAGGCGCTGG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThraspIleGIYpHeThrAlaVal 220
```

```
DB 601 ATCAGCGTCCCTTTGCAATGTCACATCTCGAGTCGAGACATGGGCTTCACTGCTGTG 660
QY 221 HisAlaLYsTYrgInSerProTPrPheGIYThrIleGIYSerValGluValTYrThrTPr 240
DB 661 CATGCCAATATCCAAAGCCGCTGGGAACTGTGACTCATCTGAAGTCACTTACTTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGIYGIYIleProTPrGluAlaTYrPheGIYArg 260
DB 721 CTTGATATGTTTCTGTGTTGATGCTGGGTGGAAATCCCATGGCAAGCACTTCTAGAGG 780
QY 261 ValIleSerSerSerSerAlaThrTYrAlaGluValIleSerPheLeuAlaIlePheGIY 280
DB 781 GTTCTCTCTTCTTCTCTGAGCACTTATCTCAAGTCTGCTCTTCTGCACTTCTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGIYAlaIleGIYAlaSerThrAspTPr 300
DB 841 TGGCTGTGATGGCCATCCAGCCCATCATCTGGGGCCATTGGAGCATCAACAGACTGG 900
QY 301 AsnGIYThrAlaTYrgIYTYrProAspProLYsThrLYsGluGluAlaAspMetIleLeu 320
DB 901 AACCACTGTCATATGGGCTTCCAGATCCCAAGACTCAAGAGAGAGCAATGATTTTA 960
QY 321 ProIleValIleGluInTYrLeuCysProValTYrIleSerPhePheGIYLeuGIYAlaVal 340
DB 961 CCAATTTGTTCTGCACTATCTCTGCTGTGATATTTCTTTCTTGTGCTTGTGTGAGTT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1021 TCTGCTGCTGTTATGATCATACAGCAATTTCTTCATCTTGTGACGAATGTCATGTTGCA 1080
QY 361 ArgAsnIleTYrgInLeuSerPheArgGluAsnAlaSerAspLYsGluIleValTPrVal 380
DB 1081 CGGAACATCTCCACACTTCTCTTCCAGCAAAATGCTTGGAGAAAGAAATCTTTGGGTT 1140
QY 381 MetArgIleThrValLeuValPheGIYAlaSerAlaThrAlaMetAlaLeuLeuThrLYs 400
DB 1141 ATGCACATCAAGCTGTTGTTGTTGAGCATTCGAAACCATGCGCTTGTGACGAA 1200
QY 401 ThrValTYrgIYLeuTYrTYrLeuSerSerAspLeuValTYrIleIleIlePheProGln 420
DB 1201 ACTGTGATGGGCTGTGGTATCCTCAGTCTGACCTTGTATCATGTTATCTTCCCCGAG 1260
QY 421 ThrValTYrgIYLeuTYrTYrLeuSerSerAspLeuValTYrIleIleIlePheProGln 440
DB 1261 CTGCTTGTGATCTCTTTGTAAAGGAAACCAACACTTATGGGCGCTGGCAGATATGTT 1320
QY 441 PheGIYLeuPheLeuArgIleThrGIYGIYGIYProTYrLeuTYrLeuGluInProLeuIle 460
DB 1321 TCTGGCTCTTCTCTGAGAAATATCTGAGGGAGGACATATCTGTATCTTCAGCCCTTGAATC 1380
QY 461 PheTYrProGIYTYrTYrSerAspLYsAsnGIYIleTYrAsnGluInArgPheProPheLYs 480
DB 1381 TTCTACCTCGGCTATTAACCTGATGATATGATATATATCAAGAAATTTCCATTATAA 1440
QY 481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTYrLeuAlaLYsTYr 500
DB 1441 ACACTTGCATGCTTACATCATCTTCAACCAACTTTCATCTCTCATCTTACCAAGATAT 1500
QY 501 LeuPheGluSerGIYThrLeuProProLYsLeuAspValPheAspAlaValAlaArg 520
DB 1501 CTAATTTGAAGAGGAACCTTGGCACTTAATTAAGATGATTTGAATGCTCTTCTTCAACA 1560
QY 521 HisSerGIYIleAsnMetAspLYsThrIleLeuValArgAsnGluAsnIleLYsLeuAsn 540
DB 1561 CACAGTGAAGAAACATGATATAGCAATTTCTTGTCAAAATAATTAATTAATTAATGAT 1620
QY 541 GluLeuAlaProValLYsProArgGluInSerLeuThrLeuSerSerThrPheThrAsnLYs 560
DB 1621 GAACCTTGCACTTGTAAACCCAGCAGAGCATGAGCCCTCAGCTCAACTTACCAATATAA 1680
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGIYSerGIYThGIYAspAsnLeuGln 580
```

Db 1661 GAGGCTTCCTGATGTTGATTCAGTCCAGAGGGTCTGGAGCTGAGATTAATTACAG 1740
RESULT 8
US-10-724-806-1
; Sequence 1, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:00805
; CURRENT APPLICATION NUMBER: US/10/724,806
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; PRIOR FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-10-724-806-1

Alignment Scores:

Pred. No.:	9.49e-280	Length:	1743
Score:	2795.00	Matches:	537
Percent Similarity:	96.6%	Conservative:	23
Best Local Similarity:	92.6%	Mismatches:	20
Query Match:	93.4%	Indels:	0
DB:	9	Gaps:	0

US-10-724-806-4 (1-580) x US-10-724-806-1 (1-1743)

Qy 1 MetProPheHisValIGlyLeuValAlaIleIleuPheThrLeuLeuIlePheLeu 20
Db 1 ATGGCTTTCATGCGAGAGAGTGAATGATATCGTTCCTTACCTTTAATTTTGGCG 60
Qy 21 ValGlyIleThrAlaAlaThrIleThrIleAsnSerGlyAsnProGluIleuSerGlu 40
Db 61 GTTGAATATGAGGCGCTGCGAGAACCAAAACAGTGCAGCGCAGAGACGCGAA 120
Qy 41 AlaIleIleValIGlyIleArgAspIleGlyLeuLeuValIGlyPheThrIleThrAla 60
Db 121 GCCATCATAGTGTGGCCGAGATATGGTTTATGGTGTGGATTACCATGACAGCT 180
Qy 61 ThrTPValIGlyIGlyIleAsnGlyThrAlaGluAlaValIleGlyProGlyCys 80
Db 181 ACCTGGCTCGAGAGAGGATATCAATGACACGCTGAGAGCGTTTATGACAGGTTAT 240
Qy 81 GlyLeuAlaTPValHisAlaProIleGlyIleSerLeuSerLeuIleLeuGlyIleu 100
Db 241 GGCTTACGCTGGGCTCAGGACCAATGGATATTCCTAGCTGATTTTACGTCGCTG 300
Qy 101 PhePheAlaIleProMetIleArgSerIleGlyIleValIleThrLeuLeuAspProHeLeu 120
Db 301 TTTCTTGGAAAAACCTATGCGTTCAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCA 360
Qy 121 IleTyrGlyIleArgMetGlyIleLeuLeuPheIleProAlaLeuMetGlyIleMetPhe 140
Db 361 ATCTATGAGAAAACGATGGGCGGAGCTCTGTTTATTCCTGACATGATGGAGAAAATGCTTC 420
Qy 141 TPValAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 421 TGGGCTGAGCAATTTCTCTGCTTGGAGCCACCATCAGCGTATCATCATGATGGAT 480
Qy 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuIleThrLeuValIGlyIle 180
Db 481 ATGCACATTTCTGTATCATCTCTGTGACCTCATTTGCCACTCTGTACACACTGGTGGAGGG 540

Qy 181 LeuTyrSerValAlaIleThrAspValAlaGluLeuPheCysIlePheIleGlyLeuTrp 200
Db 541 CTCTATTCCTGTCGCTTACACTGATGCGTTGAGCTCTTTGATTTTGTAGGGCGTGG 600
Qy 201 IleSerValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThrAlaVal 220
Db 601 ATCAGCGTCCTCCCTTGCATATTCACATCTCGAGTCGACAGATCGGGTCACTGCTGG 660
Qy 221 HisAlaIleTyrGlnSerProTrpLeuGlyThrIleGluSerValIGlyValIleThrTrp 240
Db 661 CATGCCAATACCAAAAGCCGTGGCTGGAGACTGTGACTCTCTGAACTTCACTTCTGG 720
Qy 241 LeuAspAsnPheLeuLeuMetLeuGlyIleProTrpGlnAlaIleTyrPheGlnArg 260
Db 721 CTGTATAGTTTCTGTGTTGATGCTGGGTGAATCCATGAGCAACATCTTACAGAG 780
Qy 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db 781 GTTCTCTCTTCTTCCCTCAGCCACTATGCTCAAGTGTGTCTCTCTGCGACCTTTCGGG 840
Qy 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db 841 TCCCTGTGATGGCCATCCAGCCATCTCATTTGGGCGCATTTGAGACATCAAGACTGG 900
Qy 301 AsnGlnThrAlaIleTyrGlyIleTyrProAspProIleThrIleGluAlaAspMetIleu 320
Db 901 AACCAAGCTGATATGGGCTTCACATCCCAAGACTACAGAGAGGACACATGATTTTA 960
Qy 321 ProIleValLeuGlnTyrLeuCysProValIleIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCAATTTGTTCCAGATATCTGCGCCCTGTATATTTCTTTTGTGTGTGTGTGTGTGT 1020
Qy 341 SerAlaAlaValMetSerSerAlaAspSerIleLeuSerAlaSerSerMetPheAla 360
Db 1021 TCTGCTGCTGTATATGATCAGCAATCTTCACTTGTGCGAAGTTCCAGATTTTGA 1080
Qy 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspIleGluIleValTPVal 380
Db 1081 CGGAACATCTACACACTTTCCTTCAAGCAAAATGCTTGGAGCAAAAGAAATCGTTGGCTT 1140
Qy 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400
Db 1141 ATGCAATCACAGTGTGTTGTTGTGGAGCATCTGCAGACGCATGGCTTGTGTCAGAA 1200
Qy 401 ThrValIleGlyLeuTrpIleThrLeuSerAspLeuValIleIleIlePheProGln 420
Db 1201 ACTGTGTATGGGCTGTGAGCTCAGTCTGACCTTGTATCATGATGTTATCTTCCCCAG 1260
Qy 421 LeuLeuCysValLeuPheIleIleGlyIleThrAsnThrTyrGlyAlaValAlaIleTyrIle 440
Db 1261 CTGCTTTGTGATCTTTGTTGTAAGGAACCAACACTATGGGCGCTGACAGTTATGTT 1320
Qy 441 PheGlyLeuPheLeuArgIleThrGlyGlyIleProTyrLeuTyrLeuGlnProLeuIle 460
Db 1321 TCTGGCTCTTCTCTGAGATATCTGAGGGGAGCAATCTGTATCTTACGCCCTTGATC 1380
Qy 461 PheTyrProGlyTyrTyrSerAspIleIleTyrAsnGlnArgPheProPheIle 480
Db 1381 TTTTACCTCGGTATTTACCTGATATATGATATATATATATATATATATATATATAT 1440
Qy 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaIleTyr 500
Db 1441 ACACTTGGCATGATTAACATCTTTTAACCAACATTTGATCTCTCATCTTACCAAGTAT 1500
Qy 501 LeuPheGluSerGlyThrLeuProProIleAspValPheAspAlaValAlaIleArg 520
Db 1501 CTATTTGAAAGGAAACCTTGCACCTTAATTAATTAATTAATTAATTAATTAATTAAT 1560
Qy 521 HisSerGluGluAsnMetAspIleThrIleLeuValArgAsnGluAsnIleIleLeuAsn 540
Db 1561 CACAGTGAAGAAACATGATTAAGCAATTTCTGTCAAAATTAATTAATTAATTAATTAAT 1620
Qy 541 GluLeuAlaProValIleProArgIleSerLeuThrLeuSerSerThrPheThrAsnIle 560

```
Db      1621 GAACCTTGACCTTGACACCCAGACAGACGACCCCTCAGCTCACTTCAACATATAA 1680
Qy      561 GUAUAlaLeuAspValAspSerProGluGlySerGlyThrGluAspAsnLeuGln 580
Db      1661 GAGGCTTCCTTGAGTGTGATTCAGTCCAGAGAGGCTCGGAGCTGAGATTAATTACAG 1740

RESULT 9
US-09-911-077A-9
; Sequence 9, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1813
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1761)
US-09-911-077A-9

Alignment Scores:
Pred. No.: 1,01e-279 Length: 1813
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
DB: 3 Gaps: 0

US-10-724-806-4 (1-580) x US-09-911-077A-9 (1-1813)
Qy      1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrlLeuLeuIlePheLeu 20
Db      19 ATGCTTTCAGTGTGGAAGAGCTGATAGCTATCATCGTGTCTTACCTTTAATTTTGGCG 78
Qy      21 ValGlyIleThrAlaAlaIleTyrIleAsnSerGlyAsnProGluGlnArgSerGlu 40
Db      79 GTTGAATAATGGGCTGCTCGAAGAACCAAAAACAGTGGCGACGCAAGACGCGACGCA 138
Qy      41 AlaIleIleValGluGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
Db      139 GCATCATAGTGTGGGCGGAGATATGATTTATGTTGGTGGATTTACATGACAGCT 198
Qy      61 ThrTrpValGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db      199 ACCTGGGTCGAGGAGGAGGATATCATGACAGCGTGAAGAGCTTATATACAGTTAT 258
Qy      81 GlyLeuAlaIleThrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
Db      259 GGCCTAGCTTGGGCTCAGGCAACCAATTGATATCTCTTAGCTCATTTTAACTGGCCCTG 318
Qy      101 PhePheAlaIlePheMetArgSerIleGlyTyrValIleThrMetLeuAspProPheLeuGln 120
Db      319 TTCTTTGCAAAACCTATGCTTCAAAAGGGATATGAGCCATGTATGAGCCCGTTTCAGCA 378
Qy      121 IleTyrGlyIleArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlnMetPhe 140
Db      379 ATCTATGGAAGAAACGATGGGCGGAGCTCCGTTTATTTCCGCACTGATGGGAGAAATGTT 438
Qy      141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db      439 TGGGCTGACGAAATTTCTCTGCTTTGGAGCCACCATCAGCGTATCATCATGATGGAT 498
Qy      161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrIleuValGlyGly 180

Db      499 ATGACATTTCTGTCATATCTCTGACCTCATGCTGACCTGTATACACTGGTGGAGGG 558
Qy      181 LeuTyrSerValAlaIleThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db      559 CTCTATTTCTGGGCTTACATGATATGCTTCAAGCTCTTTTGATTTTGTAGGGCTGTGG 618
Qy      201 IleSerValProPheAlaLeuSerHisProAlaValIleAspIleGlyPheThrAlaVal 220
Db      619 ATCAGCGTCCCTTTGCAATGTCATCTCAGTCAGTCAGACATCGGGGTTCATCTCTGTG 678
Qy      221 HisAlaIleTyrGlnSerProTrpLeuGlyThrIleGlySerValGluValTyrThrTrp 240
Db      679 CATGCCAAATACAAAACCGGTGGGAACTGTGACTCATCGAAGTCTACTCTTGG 738
Qy      241 LeuAspAsnPheLeuLeuMetLeuGlyIleProTrpGlnAlaIleTyrPheGlnArg 260
Db      739 CTGTATGATTTTCTGTGTGTGATGCTGGGTGAATCCATGCAAGCATATCTTACAGAG 798
Qy      261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db      799 GTTCTCTCTTCTCTCCAGCACCATATGCTCAAGTGTGTCTTCTGAGCTTTCGGG 858
Qy      281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db      859 TGCTGTGATGAGCCATCCAGCCATCTCATTTGGGGCCATTGGAGCATCAACAGACTGG 918
Qy      301 AsnGlnThrAlaTyrGlyTyrProAspProIleThrIleSerPhePheGlyLeuGlyAlaVal 320
Db      919 AACCAAGCTGATATGGGCTTCCAGATCCCAAGACTACAGAAAGAGCAGACATGATTTTA 978
Qy      321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db      979 CCAATGTTCTGCACTATCTGCCCCGTGATATTTCTTCTTGGTCTGGTGCAGT 1038
Qy      341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db      1039 TCTGCTGCTGTATATCATCATGACGATCTTCCATCTTGTGACGACGATGCTTCCATTTGCA 1098
Qy      361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrpVal 380
Db      1099 CGGAACATCTACACAGCTTCTTCCAGCAAAATGCTTGGCAAAAGAAATCGTTGGGTT 1158
Qy      381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400
Db      1159 ATGCCAAATCACAGTGTGTTGTGTGGACATCTGCAACAGCCATGCGCTTGTGACGAAA 1218
Qy      401 ThrValTyrGlyLeuTrpTyrIleSerSerAspLeuValTyrIleIleIlePheProGln 420
Db      1219 ACTGTGATGGGCTTGGTACTCTGACCTTGTATCATGCTTATCTTCCCCAG 1278
Qy      421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db      1279 CTGCTTTGTATGACTTTTGTATAGGGAACCAACCTATGGGGCGGTGGAGGTATGTT 1338
Qy      441 PheGlyLeuPheLeuArgIleThrGlyGlyGlnProTyrIleTyrLeuGlnProIleIle 460
Db      1339 TCTGGCTCTTCTCGAAGATTAACGGAAGGAGCCATATCTGTATCTTCAAGCCTTGTATC 1398
Qy      461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheIle 480
Db      1399 TTCTACCTGCTGCTATTAACCGATGATATGATATATATACAAATTTCCATTATAA 1458
Qy      481 ThrLeuSerMetValIleThrSerPheThrAsnIleCysValSerTyrLeuAlaIleTyr 500
Db      1459 ACACCTTGCATGATACATCATTTTAACCAACATTTGCATCTCTATCTGACCAAGTAT 1518
Qy      501 LeuPheGluSerGlyThrLeuProProIlePheAspValPheAspAlaValAlaIleArg 520
Db      1519 CTATTTGAAGAGGAACCTTGCACCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1578
Qy      521 HisSerGluGluAsnMetAspLysThrIleLeuValIleAsnGluAsnIleLysLeuAsn 540
```


Db 1579 CACAGTGAAGAAAACATGATTAAGACATCTTGTCAAAAATGAAAATATTAAATTAGAT 1638
QY 541 GtLeuLeuAlProValIysProArgIInsLeuThrlensSerSerThrPheThrAnlIys 560
Db 1639 GAACCTTGACCTTGTGAAGCCACAGACAGATGACCTTCACTGCTCACTTCAACATAAA 1698
QY 561 GtUAlLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAnLeuGln 580
Db 1699 GAGGCGCTTCTGATGTGATTCAGTCCAGAGAGGGGTCTGGAGCTGAAGATAATTATACAG 1758
RESULT 10
US-10-724-806-9
; Sequence 9, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARUSUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/10/724,806
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1813
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1761)
US-10-724-806-9

Alignment Scores:

Pred. No.: 1.01e-279 Length: 1813
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
DB: 9 Gaps: 0

US-10-724-806-4 (1-580) x US-10-724-806-9 (1-1813)

QY 1 MecPProPheHisValGluGlyLeuValAlaIleIleLeuPheThrlensLeuIlePheLeu 20
Db 19 ATGGCTTTTCATGTGAGAGCATGATGATCATCTGTTCTTAATTTTGGCTG 78
QY 21 ValGlyIleThrAlaIleThrIleAsnSerGlyAsnProGluGluArgSerGlu 40
Db 79 GTTGGATATGGGCTGCTGAGAACCAAAACAGTGGCAGCCGACAGAGGCCGAGGAA 138
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
Db 139 GCATCATATGTTGGTGGCCGAGATATTGTTATTGGTTGGTATTCATGACAGCT 198
QY 61 ThrTrpValGlyGlyGlyTrpIleAsnGlyThrAlaGluAlaValIleGlyProGlyCys 80
Db 199 ACCTGGGTGGAGAGGATATCAATGCGACAGCTGAACAGTTATATGTACAGGTTAT 258
QY 81 GtLeuLeuAlaTrpAlaHisAlaProIleGlyTrpSerIleuSerIleuIleuGlyGlyLeu 100
Db 259 GGGCTTGGCTTGGGCTCAGGACCAATTTGATATCTTCTTATGATTTTAAAGTGGCTG 318
QY 101 PhePheAlaIysProMetArgSerIleGlyTrpValThrMetLeuAspProPheIysGln 120
Db 319 TTTCTTGGCAAAACCTATGCGTTCAAGGGGTATGTGACATGTTAGACCCGTTTCAGCAA 378
QY 121 IleTrpGlyValArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
Db 379 ATCTATGAAAGAACGATGGGCGGACTCTCTTTATTCTCTCACTGATGAGGAGAAATGTTTC 438

QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 439 TGGGCTGCAAGCAATTTTCTCTGCTTGGGAGCCACCATGACGCTGATCATCATGTGAT 498
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTrpThrLeuValGlyGly 180
Db 499 ATGCACATTTCTGTCATCATCTCTGACATCTGACCTGTCACACTGTGAGGGAGGG 558
QY 181 LeuTrpSerValAlaTrpThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db 559 CTCTATTCTGGGCTTACACTGATATGTCGTTCAGCTCTTTTGCAATTTTGGAGGCTGTG 618
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db 619 ATCAGGCTCCCTTTCATATGTCACATCTGACAGTGGCAGACATCGGGTTCACTGCTGTG 678
QY 221 HisAlaIleGlyTrpGlnSerProTrpIleuGlyThrIleGluSerValGluValIleTrpTrp 240
Db 679 CATGCCAAATACCAAAAGCCGTGGCTGGGAACGTGTGACTCATCTGAAAGCTTCTTGG 738
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTrpPheGlnArg 260
Db 739 CTGTATAGTTTCTTCTTCCACAGCACCTAATGCTCAAGTCTGTCTTCTGCGAGCTTGGG 858
QY 261 ValLeuSerSerSerSerAlaThrTrpAlaGlnValLeuSerPheLeuAlaAlaPheGly 280
Db 799 GTTCTCTCTTCTTCCACAGCACCTAATGCTCAAGTCTGTCTTCTGCGAGCTTGGG 858
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db 859 TGCCTGTGATGGCCATCCAGCCATCATCTATGGGGCCATTTGGAGATCAACAGCTGG 918
QY 301 AsnGlnThrAlaTrpGlyTrpProAspProIysThrIleGluGlnAlaAspMetIleLeu 320
Db 919 AACCAAGCTGCAATGAGGCTTCCAGATCCCAACACTCACAAAGGCGACAGATGATTTTA 978
QY 321 ProIleValLeuGlnTrpLeuCysProValTrpIleSerPhePheGlyLeuGlyAlaVal 340
Db 979 CCAATGTTCTGTCAGATCTCTGCTCGGTGATATTTCTTGGTCTTGGTGAGCTT 1038
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1039 TCTGCTGCTGTTATGTCATCAGCAGATCTTCCATCTTGTCAACAAGTTCCATGTTTGA 1098
QY 361 ArgAsnIleTrpGlnLeuSerPheArgGlnAsnAlaSerAspIysGlnIleValIleTrpVal 380
Db 1099 CGAAGATCTACCAAGCTTCTCTTCAACAAATGCTTGGACAAAGAAATCGTTGGGTT 1158
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIys 400
Db 1159 ATCGAATCACAAGTGTGTTGTTGGAGATCTGCAACGCCATGGCTTGCAGAGAAA 1218
QY 401 ThrValIleGlyLeuTrpTrpIleuSerSerAspLeuValTrpIleIleIlePheProGln 420
Db 1219 ACTGTATAGGGCTCGGTGACCTCAGTTCGACCTGTTTACATCGTTATCTTCCCCCAG 1278
QY 421 LeuLeuCysValLeuPheIleIysGlyThrAsnThrTrpGlyAlaValAlaGlyTrpIle 440
Db 1279 CTGCTTGTGATCTTGTGTAAGGAAACCAACCACTATGGGCGGGAGGTTATGTT 1338
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTrpIleuTrpLeuGlnProLeuIle 460
Db 1339 TCTGGCTCTCTCTGGAATTAACCTGAGGGGAGCCATATCTGATCTTTCAGCCCTGATC 1398
QY 461 PheTrpProGlyTrpTrpSerAspIysAsnGlyIleTrpAsnGlnArgPheProPheIys 480
Db 1399 TTTCTACCTGGCTATTAACCTGATGATATGATATATATATGAAATTTTCAATTTAAA 1458
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTrpLeuAlaIysTrp 500
Db 1459 ACACTTGCATGTGATCATCTTCAACCAATTTGATCTCTCTATCTAGCCAACTAT 1518
QY 501 LeuPheGluSerGlyThrLeuProProIysLeuAspValPheAspAlaValAlaIleArg 520


```
Db      1519 CTATTGTAAGAGGAACTTCGACCTTAATTAAGATTGATGCTGTTGTCACGA 1578
Qy      521 HisSerGluGluAspMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
Db      1579 CACAGTGAAGAAAATATGATAGACAAATCTCTGCAAAAATGAAAATTTAAATTAGAT 1638
Qy      541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db      1639 GAACCTTGACCTTGTAAGCCACGACAGACGATGACCTCAGCTCAACTTCAACATATAA 1698
Qy      561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
Db      1699 GAGGCTTCCTTGATGATTGATTCAGTCCAGAGAGGCTGGGACGTGAAATTAATTACAG 1758

RESULT 11
US-10-241-784-1
; Sequence 1, Application US/10241784
; Publication No. US20040048261A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Corporation
; TITLE OF INVENTION: Invertebrate Choline Transporter Nucleic Acid, Polypeptides and U
; FILE REFERENCE: MO7218
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 1833
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1833)
; OTHER INFORMATION:
US-10-241-784-1

Alignment Scores:
Pred. No.: 4,93e-144 Length: 1833
Score: 1491.50 Matches: 294
Percent Similarity: 72.7% Conservative: 81
Best Local Similarity: 57.0% Mismatches: 122
Query Match: 49.8% Indels: 19
DB: Gaps: 7

US-10-724-806-4 (1-580) x US-10-241-784-1 (1-1833)
Qy      4 HisValGluGlyLeuValAlaIleIleLeuPheTyrrLeuLeuIlePheLeuValGlyIle 23
Db      7 AATATCGCTGGCGGTGAGCATCGTCTTCTTAACCTTCACTCGGTCTGCGCATTT 66
Qy      24 TrpAlaAlaTrpLysThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIle 43
Db      67 TGGGCGGCGCGAAG---AAGCAGTCCGGCATGATTCGAG-----GAGAGAGCGATG 117
Qy      44 ValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAlaThrTrpVal 63
Db      118 CTGGCGGCGAGCTCCATCGGCTCTTCGTGGGCATCTTACCATGACGGCCACCTGGGCTG 177
Qy      64 GlyGlyGlyTyrrIleAsnGlyThrAlaGluAlaValTyrrGlyProGlyCysGlyLeuAla 83
Db      178 GGTGGCGGCTACATCAACGCGACGCGGCGAGGCTATATAC-----ACATCGGAGCTGGTG 231
Qy      84 TrpAlaHisAlaProIleGlyTyrrSerLeuSerLeuIleLeuGlyGlyLeuPhePheAla 103
Db      232 TGGTGCCAGGCTCCATTTGGATACGCTTAAGCTTGATTTGGTGGCATCTTCTTTGGC 291
Qy      104 LysProMetArgSerLysGlyTyrrValThrMetLeuAspProPheLysGlnIleTyrrGly 123
Db      292 AATCCCATCGGACGAGGAGGTTACATCACCATGTTGATCCGTTGACGAGATTCCTTTGGT 351
Qy      124 LysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLeuPheThrAlaAla 143
```

```
Db      352 GAGCGGATGGAGAGATTCCTCTCTGCGCTTATCGCGGAGGCTTTTGGCACCC 411
Qy      144 AlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAspValAsnIle 163
Db      412 GGCATCTCGGCTGACCTTGCGGCCACTATCGGATGATCGAATGATACCGCACCC 471
Qy      164 SerValIleValSerAlaLeuIleAlaIleLeuTyrrThrLeuValGlyGlyLeuTyrrSer 183
Db      472 TCGGATGATCCGTCTCTCGATCGCATCTTCAACACACTGTTGGGAGCTGATACCC 531
Qy      184 ValAlaTyrrTrpAspValValGluLeuPheCysIlePheIleGlyLeuTrpIleSerVal 203
Db      532 GTGGCGTATACGAGAGTATATCAGTTGTTCTGATCTTCAATCGGCTGTGATGTCATTT 591
Qy      204 ProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaValHisAlaLys 223
Db      592 CCCTTCGCTGGAGGAGAACGAGAGCTGGGCGAGCCG-----AGTGAC 633
Qy      224 TyrrGlnSerProTrpLeuGlyThrIleGlnSerValGluValTyrrThrTrpLeuAspAsn 243
Db      634 CTGGAGGTGATGTGATTTGGGACGACGCTGAAGGCTTAAGAAAGCATTTGCTGTACTATAC 693
Qy      244 PheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrrPheGlnArgValLeuSer 263
Db      694 GCGTGTGCTGCTGCTTTTGGTGGCATTCCTTGGCAGGTCTACTTCCAGCG----- 744
Qy      264 SerSerSerAlaThrTyrrAlaGlnValLeuSerPheLeuAlaAlaPheGlyCysLeuVal 283
Db      745 ---CAAAACGGCAGGAGAGGCCCGACCTTCGCTATGTTGACGCGCGGATGCATTTTG 801
Qy      284 MetAlaLeuProAlaIleCysIleGlyValAlaIleGlyAlaSerThrAspTrpAsnGlnThr 303
Db      802 ATGGCATATCCCGCCGCTGCTCATCGAGCATTTGCCAAGCTACACCTTGAACACAGACA 861
Qy      304 AlaTyrrGlyTyrrProAspProLysThrLysGluGluAlaAspMetIleLeuProIleVal 323
Db      862 GATTACAAAGGACCTTATCCCTGACCGTGACGACGACGACGATATCTGCGCCATGGTG 921
Qy      324 LeuGlnTyrrLeuCyProValTyrrIleSerPhePheGlyLeuGlyAlaValSerAlaAla 343
Db      922 CTGCAGTACCTACGCGTGACTGTGTCCTTCTTGGATTTGGGCGCTTCCCGCGCC 981
Qy      344 ValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIle 363
Db      982 GTGATGTCTCTCGCCGACCTCTCGGTGCTCTCCCGCGCTCCATGTTGCTGGAAACGGG 1041
Qy      364 TyrrGlnLeuSerPheArgGluAsnAlaSerAspLysGluIleValTrpValMetArgIle 383
Db      1042 TACAAATTTGATTTTCCGTACAGAGCGTCCGAGATGAAATATTTGGGTATGCGAGTTC 1101
Qy      384 ThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLysThrValTyrr 403
Db      1102 GCCATCATTTGTGTGGGATCTCGCTACCATCATGAGCCCTCACCATTCCTCCATCTAC 1161
Qy      404 GlyLeuTrpTyrrLeuSerSerAspLeuValTyrrIleIleIlePheProGlnLeuLeuCys 423
Db      1162 GGTTTGTGTGTCATGCTGCGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1221
Qy      424 ValLeu---PheIleLysGlyThrAsnThrTyrrGlyAlaValAlaIleGlyTyrrIlePheGly 442
Db      1222 GTGGTCACTTCAAGAACGACCTGACACACGTCAGGACGCTTCCGATACATTTGGTGGC 1281
Qy      443 LeuPheLeuArgIleThrGlyGlyGluProTyrrLeuGlnProLeuIlePheTyrr 462
Db      1282 CTGGCATCTCCGATCTGCGGCGGTGAGGCGCATCTTGGGACCTGCGCATTTGATCAAGTAT 1341
Qy      463 ProGlyTyrrTyrrSerAspLysAsnGlyIleTyrrAsnGlnArgPheProPheLysThrLeu 482
Db      1342 CCGGCTACGACGAGAGAACCAAG-----GAGCAGATGTTCCCTTCCGACCAAG 1392
Qy      483 SerMetValThrSerPhePheThrAsnIleCysValSerTyrrLeuAlaTyrrLeuPhe 502
```

Db 1393 GCCAGCTGCTCAGCCTGCTACGCTCATCTCGGTCTCTGTGTGAGCTAAAGATGTTT 1452
Qy 503 G|USeRg|YThrLeuProProlYsLeuAspValPheAspAlaVal 518
Db 1453 GAGTCGGGCAAGTTGCCGCCAGCTACGACTTCCGCTGTGGTC 1500
RESULT 12
US-09-911-077A-7
; Sequence 7, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: AAPSUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)..(1739)
US-09-911-077A-7
Alignment Scores:
Pred. No.: 4,34e-139 Length: 1985
Score: 1444.50 Matches: 292
Percent Similarity: 68.6% Conservative: 93
Best Local Similarity: 52.0% Mismatches: 145
Query Match: 48.3% Indels: 31
Gaps: 9
US-10-724-806-4 (1-580) x US-09-911-077A-7 (1-1985)
Qy 7 G|YLeuValAlaIleIleuPheTYrLeuLeuIlePheLeuValGlyIleTrrAlaIa 26
Db 24 GGNATGTCGGCCATGTTGTTCTTCTACGCTCATCTTGTGCGTAATATGGGGGT 83
Qy 27 T|P|YsThrLysAsnSerGlyAsnProGluGluArg-----SerGluAlaIle 42
Db 84 AGAAATTCGAAAGTTCGAAAGAGCTTGAATCGAAGCCGCCGCCGCGAGAGAGGTG 143
Qy 43 I|e|ValG|Y|A|G|A|S|P|I|e|G|Y|LeuValG|Y|P|H|e|ThrAlaI|H|T|P| 62
Db 144 ATGTTAGCTGGGAGAAACATCGAAGCTTGTGCGAATTTTCACATGACTGCCAGTGG 203
Qy 63 V|A|G|Y|G|Y|Y|Y|I|L|e|S|N|G|Y|ThrAlaG|U|A|V|A|I|Y|G|Y|P|R|O|G|Y|C|Y|S|G|Y|L|E|U| 82
Db 204 GTTGGGGGTCTTATATCATGGAACGCCGAGCTCTGTATATGGA-----GGTCTC 257
Qy 83 A|L|A|T|P|A|H|I|A|P|R|O|I|E|G|Y|Y|S|E|R|L|E|U|I|L|E|U|G|Y|G|I|Y|L|E|U|P|H|E| 102
Db 258 CTTGGATGTCAGGCTCAGTGGATATGCAATTTCCCTGTATGGAGAGACTATTTTC 317
Qy 103 A|A|L|Y|S|P|R|O|M|E|C|A|R|G|S|E|R|Y|S|G|Y|Y|V|A|I|T|H|E|T|L|E|U|A|S|P|R|O|P|H|E|S|G|I|N|I|E|Y| 122
Db 318 GCAAGAAATGCGAGAGAGAGATATATTAACAATCTCGATCTTTTTCAGACAATAT 377
Qy 123 G|Y|Y|S|A|R|G|M|E|G|Y|G|Y|L|E|U|E|U|P|H|E|I|P|R|O|A|L|E|U|M|E|G|Y|G|U|L|E|T|P|H|E|T|P|A|I|A 142
Db 378 GGCACACGATCGTGGCTTGATGTTCAGCACTTCTTGTAACATTTCTGACA 437
Qy 143 A|A|A|A|I|I|P|H|E|S|E|R|A|L|E|U|G|Y|A|I|A|T|H|I|L|E|S|E|R|V|A|I|I|E|L|E|A|S|P|V|A|S|P|A|S|N 162
Db 438 GAGGCATCTTTTCCGCACTTGTGCAACACTCTCGTATATCTTGGAAATCGACATGAT 497
Qy 163 I|L|E|S|E|R|V|A|I|L|E|V|A|S|E|R|A|L|E|U|I|L|E|U|I|L|E|U|Y|Y|T|H|R|L|E|U|V|A|G|Y|G|Y|L|E|U|Y| 182

Db 498 GCATCAGTACCCCTGTGCGCTGTATTCGCTATTCACATTCACCGGTGATCTAT 557
Qy 183 S|E|R|V|A|I|A|T|Y|T|H|A|S|P|V|A|I|G|I|N|L|E|U|P|H|E|C|Y|S|I|L|E|P|H|I|L|E|G|Y|L|E|U|T|P|I|L|E|S|E|R| 202
Db 558 GCAATGCGCTACACTGACGCTGTTCACTAATTTTGTGCTGCGTTTGTGGGTTTGC 617
Qy 203 V|A|I|P|R|O|P|H|E|I|L|E|U|S|E|R|H|I|S|P|R|O|A|V|A|I|T|H|A|S|P|I|L|E|G|Y|P|H|E|T|H|A|V|A|H|I|S|A|I|A 222
Db 618 GTGCGCGCGCTATGTCATGATGATGGTGGCAAGATATTTCCAGAAATGCA-----668
Qy 223 L|Y|S|Y|Y|G|I|N|S|E|R|P|R|O|T|R|E|U|G|Y|T|H|I|L|E|G|U|S|E|R|V|A|I|-----G|U|V|A|I|Y|T|H|T|P|L|E|U| 241
Db 669 -----GGCAGCTGATTTGAGAGATTGGAGATTGCAAGAAACATCTCTGGAAT 719
Qy 242 A|S|P|A|S|N|P|H|E|L|E|U|L|E|U|M|E|T|L|E|U|G|Y|Y|I|P|R|O|T|R|G|I|A|L|A|T|P|H|E|G|I|N|A|R|G|V|A|I| 261
Db 720 GATTCATGCTCTTCTCTTGTGAGGAAATTCATGGCAAGTGTACTTCGCAAGAGAT 779
Qy 262 L|E|U|S|E|R|S|E|R|S|E|R|A|T|H|Y|R|A|G|I|N|V|A|L|L|E|U|S|E|R|P|H|E|U|A|I|A|P|H|E|G|Y|C|Y|S| 281
Db 780 CTCTCTCAAAACACTGCTCATGGAGACAGAGCTTGTCTGTTGTGGCGGCGCTGGATGC 839
Qy 282 L|E|U|V|A|L|E|R|A|L|E|U|P|R|O|A|I|L|E|C|Y|S|I|L|E|G|Y|A|I|L|E|G|Y|A|L|S|E|R|T|H|A|S|P|T|P|A|S|N 301
Db 840 ATTCTCATGGCATTCACACAGCGTTGATCGGCAATTGCCAGAAACAGACTGGAGA 899
Qy 302 G|I|N|T|H|A|I|T|Y|R|G|Y|R|P|R|O|A|S|P|R|O|L|Y|S|T|H|Y|S|G|I|U|G|I|A|A|-----316
Db 900 ATGACTGATATTTCCCATGAAACAAATGAGACTAGAGTGCATTCGATTCACCGGATAG 959
Qy 317 ---A|S|P|M|E|I|L|E|U|P|R|O|I|L|E|U|G|I|N|Y|L|E|C|Y|S|P|R|O|V|A|I|Y|I|L|E|S|E|R|P|H|E| 335
Db 960 AGAAATATGTTGGTCCCGTTGGATTCAGTATCTTACGCCAAGATGGGCGCTTAT 1019
Qy 336 G|Y|L|E|U|G|Y|A|V|A|S|E|R|A|A|A|V|A|L|M|E|T|S|E|R|S|E|R|A|A|S|P|S|E|R|I|L|E|U|S|E|R|A|I|A 355
Db 1020 GGACTCGGCGAGTGCCTGCTGTAAATGTCATCGCAGATTCACTGACTATACGCA 1079
Qy 356 S|E|R|S|E|R|P|H|E|A|A|R|G|A|S|N|I|L|E|Y|G|I|N|L|E|U|S|E|R|P|H|A|R|G|I|N|S|A|L|S|E|R|A|P|L|Y| 375
Db 1080 GCATCATATTTGCTCACAACATCTGAAAGCTCACAATTCGCCCTCAGCGGTGAAAA 1139
Qy 376 G|U|I|L|E|V|A|I|T|P|V|A|L|E|R|A|G|I|L|E|T|H|V|A|L|E|U|V|A|P|H|E|G|Y|A|L|S|E|R|A|A|T|H|A|R|A|L|A|E|R| 395
Db 1140 GAAATCATATTTGATGATGATAGATAGCCATCTGTGTTGATATCTGGCAACATCATG 1199
Qy 396 A|A|L|E|U|L|E|U|T|H|Y|S|T|H|V|A|I|Y|G|Y|L|E|U|T|P|Y|R|L|E|U|S|E|R|S|E|R|A|S|P|L|E|U|V|A|I|Y|I|L|E| 415
Db 1200 GCACTTACCATTCATTCATCTATGGGCTTGGTATCTTTGCAAGATTTGGCTACGTC 1259
Qy 416 I|L|E|I|P|H|E|P|R|O|G|I|N|L|E|U|C|Y|S|V|A|L|E|U|P|H|E|I|L|E|Y|G|Y|T|H|Y|S|N|H|Y|R|G|Y|A|I|A 435
Db 1260 ATACTTCTCCCAACTATATATGTGTATATATGCCAGTAGCAATATAGTAGCTCA 1319
Qy 436 V|A|I|A|G|Y|Y|I|L|E|P|H|E|G|Y|L|E|U|P|H|E|U|A|G|I|L|E|T|H|G|Y|G|Y|L|E|U|P|R|O|Y|R|L|E|U|Y| 455
Db 1320 TTGGCTGCTATGCACTGCGTCTTGTGCTCCGTTGATTTGGAGCGACACTTGTATCG 1379
Qy 456 L|E|U|G|I|N|P|R|O|L|E|U|I|L|E|P|H|E|Y|R|P|R|O|G|Y|Y|Y|R|S|E|R|A|S|P|L|Y|S|N|G|Y|I|L|E|Y|R|A|E|N|G|I|N 475
Db 1380 CTGCGACGCTTCTTCATATCAATGATAT-----ACGAGATGGGGTA-----CAG 1424
Qy 476 A|R|G|P|H|E|R|O|P|H|E|L|Y|S|T|H|R|L|E|U|S|E|R|M|E|T|V|A|I|T|H|S|E|R|P|H|E|T|H|Y|S|N|I|L|E|C|Y|S|V|A|L|S|E|R| 495
Db 1425 TATTTTCCATTCAGGCACTGCTATGTATCTTCAATGCTACTATCTTACATTTGATCA 1484
Qy 496 T|Y|R|E|U|A|L|Y|S|T|Y|R|L|E|U|P|H|E|G|U|S|E|R|G|Y|T|H|L|E|U|P|R|O|L|Y|S|L|E|U|A|S|P|V|A|P|H|E|A|S|P 515
Db 1485 ATACATCGAGAGAGTGTCAATCGGAGCGTTGTCTCCGAGATGGACGTAAGGAT 1544
Qy 516 A|A|V|A|I|V|A|I|-----A|A|R|G|H|S|E|R|G|I|U|G|I|N|S|E|R|A|S|P|L|Y|S|H|I|L|E|U|V|A|I| 532
Db 1545 TGTGTAGTAATATTCGATAGATCATGTAACCTTCCTCGCTAGATGATTCGTTGCTGTT 1604

QY 533 ArgAnGIuaSnIleYleu-----AsnGIuLeuAlProValIysProArg 548
Db 1605 AGTAGTGAGACCTTGAAATATGMAAGCTCCAAACGGAACCGGCTCCAGTACATCCGAAC 1664
QY 549 Gln 549
Db 1665 CAA 1667
RESULT 13
US-10-724-806-7
Sequence 7, Application US/10724806
Publication No. US20040248838A1
GENERAL INFORMATION:
APPLICANT: BLAKELY, RANDY D.
APPLICANT: APPARUNDARAM, SUBRAMANIAM
APPLICANT: FERGUSON, SHAWN
TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
FILE REFERENCE: VBLT:008US
CURRENT APPLICATION NUMBER: US/10/724,806
CURRENT FILING DATE: 2003-12-01
PRIOR APPLICATION NUMBER: US/09/911,077A
PRIOR FILING DATE: 2001-07-23
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 1985
TYPE: DNA
ORGANISM: Caenorhabditis elegans
FEATURE:
NAME/KEY: CDS
LOCATION: (9)..(1739)
US-10-724-806-7
Alignment Scores:
Pred. No.: 4,346-139 Length: 1985
Score: 1444.50 Matches: 282
Percent Similarity: 68.6% Conservative: 93
Best Local Similarity: 52.0% Mismatches: 145
Query Match: 48.3% Indels: 31
Gaps: 9
US-10-724-806-4 (1-580) x US-10-724-806-7 (1-1985)
QY 7 GlyLeuValAlAlIleLeuPheTyLeuLeuIlePheLeuValGIYIleTrrAlaLa 26
Db 24 GGTAATCGTGCCATGTTCTTACGCGTCATCTTGTAATATGAGCGGAT 83
QY 27 TrrIysThrIysAsnSerGIYAsnProGIuIArg-----SerGIuAlaIle 42
Db 84 AGAAATCGAAAAGTTCAAAAGAGCTTGATCAGAAAGCCGCGCGCGAGAGAGGTG 143
QY 43 IleValGIYGIYArgAspIleGIYLeuLeuValGIYGIYLeuPheThrMetThrAlaTrrTP 62
Db 144 ATGTAGCTGGGAGAAACATCGGAATCTTGTCCGAATTTTCAACAATGATCGCAAGTGG 203
QY 63 ValGIYGIYGIYTrIleAsnGIYThrAlaGIuAlaValTrGIYProGIYCySGIYLeu 82
Db 204 GTTGGCGGTGCTTATATCAATGAAACCGCGGAGGCTCTGTATTAATGA-----GGTCTC 257
QY 83 AlaTrrAlaHisAlaProIleGIYTrSerLeuSerLeuIleLeuGIYGIYLeuPhePhe 102
Db 258 CTGGATGTCAAGCTCCAGTGTGATATGCAATTTCCCTGTATAGGAGACTACTTTTC 317
QY 103 AlAlysProMetArgSerIysGIYTrValIThrMetLeuAspProPheIysGlnIleTr 122
Db 318 GCAAGAGAAATTCGAGAGAGATATATACATGCTCGATCTTTTACAGCAAAATAT 377
QY 123 GIYIysArgMetGIYGIYLeuLeuPheIleProAlaLeuMetGIYGIYLeuPheTrrAla 142
Db 378 GGCCCAACGAATCGTGCTGTGATGTATGTCAGCACTTGTGGTGAACATTTCTGGACA 437
QY 143 AlaAlaIlePheSerAlaLeuGIYAlaThrIleSerValIleIleAspValAspValAsn 162

Db 438 GCAGCATTTCTTGGGCACTTGTGCAACACTGTGCAATTTCTGGAATGACATGTAAT 497
QY 163 IleSerValIleValSerAlaLeuIleAlaIleLeuTrIleLeuValGIYGIYLeuTr 182
Db 498 GCATCAGTACCCCTGTCCGCTGTATCCGATTTCTACATTCACCGGTGGATACAT 557
QY 183 SerValAlaTrIThrAspValValGIuLeuPheCySIIlePheIleGIYLeuTrrIleSer 202
Db 558 GCAGTCGGGTACACGACGATGCTTCACTATTTGCAATTTCTGTGCGTTGTGGGTTC 617
QY 203 ValProPheAlaLeuSerHisProAlaValIThrAspIleGIYLeuPheThrAlaValHisAla 222
Db 618 GTGCCGGCGCTATGTGTCATGATGCTCCGAAGATATTTCCAGGAATGCA----- 668
QY 223 LysTrGIuIserProTrrLeuGIYTrIleGIuSerVal---GIuValTrIThrTrrLeu 241
Db 669 -----GGGCACTGATTTGGAGATTGGAGATTGGAAGATTCAAAGAAACATCTCTGGAAT 719
QY 242 AspAsnPheLeuLeuLeuMetLeuGIYGIYIleProTrrGIuAlaTrIlePheGIuArgVal 261
Db 720 GATTGCATGCTTCTCTGCTTGTGAGAGAAATTCATGCGCAAGTGTACTTCCAAAGAGTT 779
QY 262 LeuSerSerSerSerAlaThrTrIValGIuValLeuSerPheLeuAlaAlaPheGIYCyS 281
Db 780 CTCTCTCAAAAACGCTGCTCATGAGACACAGACGTTGTGTGTGTGCGCGCGCTCGGATC 839
QY 282 LeuValMetAlaLeuProAlaIleCySIIleGIYAlaIleGIYAlaSerThrAspTrrAsn 301
Db 840 ATTCTCATGGGATTCACCAAGCGTTGATCGGTGCAATTTGCCAGGAACACAGACTGGAGA 899
QY 302 GlnThrAlaTrGIYTrIProAspProIysThrIysGIuIuAla----- 316
Db 900 ATGACTGATTTATTCCTCCATGGAACATGGAACATGAACTAAGTCATTCACCGGATAG 959
QY 317 ---AspMetIleLeuProIleValLeuGIYTrIleCySProValTrIleSerPhePhe 335
Db 960 AGAACATAGTGGTGTCCGTGGTATTCAGATTCATTCAGCCAAAGTGGGTCCCTTTATT 1019
QY 336 GIYLeuGIYAlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAla 355
Db 1020 GGACTCGGCGAGTGTGCGGTCTGTAAATGATCATGCAAGATTCATCTGACTATACAGA 1079
QY 356 SerSerMetPheAlaArgAsnIleTrGIuLeuSerPheIleArgGlnAsnAlaSerIys 375
Db 1080 GCATCAATGTTTGTCTCACAAATCTGGAAGCTCAAAATTCGCTCAGCGCTGTAATAA 1139
QY 376 GIuIleValIThrValMetArgIleThrValLeuValPheGIYAlaSerAlaThrAlaMet 395
Db 1140 GAAATGATTAATTTGATAGATGAAATACCATCATCTGTGTGGTATCATGGCAACATCATG 1199
QY 396 AlaLeuLeuThrIysThrValTrGIYLeuTrIYLeuSerSerAspIleValTrIle 415
Db 1200 GCATTCACCATTCATTCATCATGAGCTTGGTATCTTTGTCAGATTTGGTCTACGTC 1259
QY 416 IleIlePheProGlnLeuLeuCySValLeuPheIleIysGIYThrAsnThrTrGIYAla 435
Db 1260 ATACTCTTCCCTCACTATATGTGTGTATATATAGCAACGTAGCAAAACGATAGGCTCA 1319
QY 436 ValAlaGIYTrIlePheGIYLeuPheLeuArgIleThrGIYGIYGIYProTrIYLeu 455
Db 1320 TTGGCTGGCTATGCGTGGTGTGTGCTCCGTTGATGGAGGAGGACCTGTATTCG 1379
QY 456 LeuGlnProLeuIlePheTrIProGIYTrIYTrSerAspIysAsnGIYIleTrIYAsnGln 475
Db 1380 CTGGCAGCGTCTTCATTTATCAATGTAT-----ACGATGGGGTA-----CAG 1424
QY 476 ArgPheProPheIysThrLeuSerMetValIThrSerPhePheThrAsnIleYsValSer 495
Db 1425 TATTTCCATTCACGACACATGCTATATCTTCAATGGCTACATCATCATTTATCA 1484
QY 496 TrIleuAlaIysTrIleuPheGIuSerGIYThrIleuProIYLeuAspValIleAsp 515

Db 1485 ATACATCGAGAGCTGTTCAATCGGACGTTTGTCTCCGAGTGGACGTAATGGT 1544
Qy 516 AlavaVal-----AlaargHisSerGluGluAsnMetAspLysThrIleLeuVal 532
Db 1545 TGTGTAGTAAATATTCGATAGATCATGCCCTTCCTCAATGTATGCTTGTGT 1604
Qy 533 ArgAsnGluAsnIleLysLeu-----AsnGluLeuAlaProValLysProArg 548
Db 1605 AGTAGAGAACCTTGATATGTGAAGCTCCAAACGAAACCCGGCTCAGTACATCCGAAC 1664
Qy 549 Gln 549
Db 1665 CAA 1667
RESULT 14
US-11-097-143-41593/c
; Sequence 41593, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41593
; LENGTH: 4223
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-41593
Alignment Scores:
Pred. No.: 2,57e-135 Length: 4223
Score: 1413.00 Matches: 305
Percent Similarity: 59.9% Conservative: 80
Best Local Similarity: 47.4% Mismatches: 116
Query Match: 47.2% Indels: 143
DB: 13 Gaps: 10
US-10-724-806-4 (1-580) x US-11-097-143-41593 (1-4223)
Qy 4 HisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeuValGlyIle 23
Db 3217 AATATCGCTGGCGGTGGACATCGTCTTCTACTCTGTATCCGTGCTGGCAT 3158
Qy 24 TTPAlaIaIaTrpLysThrLysAsnSerGlyAsnProGluGluArgSerGlyAlaIleIle 43
Db 3157 TGGGCGCGTGGCAG--AAGCAGTCCGGCAATGATTCGAG-----GAGGAGTTCATG 3107
Qy 44 ValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAlaTrpVal 63
Db 3106 CTGGCCGAGCGCTCCATCGGCTCTTGTGGGCACTTCCACCATGACGGCACCCTGGG 3047

Qy 64 GlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAla 83
Db 3046 GGTGGCGGCTATCATCAACGAGCGGCGAGGATATATAC-----ACATCGGCTTGGTG 2993
Qy 84 TTPAlaIaIaProIleGlyTyrSerLeuSerLeuIle----- 96
Db 2992 TGGTGGCAGGCTCCATTTGGATATACGCTTAAGCTGTGTATTGGTAATCACTTACAG 2933
Qy 97 -----LeuGly 98
Db 2932 ATTCCAAATTAACTTCTTTTAAACAAAAGATTTCTATCTCTTCCAAACTTAACT 2873
Qy 99 GlyLeuPhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPhe 118
Db 2872 GGCATCTTCTTGGCCAAATCCCATGCGCAAGCAGGTTATCATCACATGTTGATCCGTTG 2813
Qy 119 LysGlnIleTyrGlyArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlu 138
Db 2812 CAGGATTCCTTGGTGGAGGATGGAGATTTGCTCTTCCGCGCTTATCGGTTAG 2753
Qy 139 MetPheTrpAlaAlaIaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp 158
Db 2752 GTCTTTGGGACCGCGCATCTGCGTGCATTTGGCGGCACCTTATCGGTATCATCGAC 2693
Qy 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuVal 178
Db 2692 ATGATCACCGCACCTCGGATCGTATGATCTGCTCGTGCATGCGCATCTTCACACATGTTG 2633
Qy 179 GlyGlyLeuTyrSerValAlaTyrThrAspValAlaGlnLeuPheCysIlePheIleGly 198
Db 2632 GGTGGACTGATCTCCGTGGGTATGAGCATGTATCAAGTTGTTCTGCACTTCATCGGT 2573
Qy 199 LeuThrIleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThr 218
Db 2572 CTGTGATGTGCATCTCCCTTCCCTGAGCAAGACACATGCGGACGCTG----- 2522
Qy 219 AlavaIaIaLysTyrGlnSerProTyrLeuGlyThrIleGluSerValGluValTyr 238
Db 2521 -----AGTACCTGGAGGTGGATGGATTGGGACGCTGAGCCTTAAAGCATTTGG 2471
Qy 239 ThrTrpLeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPhe 258
Db 2470 CTGTACATGACATCAACCGCTTGTGCTGCTGTTGGTGGAATTCCTGGCGGTCTACTTC 2411
Qy 259 GlnArgValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIa 278
Db 2410 CAGCGGAGTGTGCTCCAGCAAAACGGCAGGAAGGCCCAAGCTTCTGTCTTATGTTGGAGCC 2351
Qy 279 PheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAla----- 296
Db 2350 GCCGATGATTTGTATGGCATTTCCCGGCTCATCGAGCATGTCGAAGGCTTACA 2291
Qy 296 ----- 296
Db 2290 CGTAGTCTTAGTAGACGACACTAGCAATTAATAACAGATATCCCAATAATAGAAAGC 2231
Qy 296 ----- 296
Db 2230 ACATATCTTGGTGGGGGCGAGATTACTATGACCTTTGACTATATGCTATTTTCTT 2171
Qy 297 ---Ser---ThrAspTrpAsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGlu 315
Db 2170 TTCTGTACAGCTTGGAAACGACACATATTCAGAGGACCCTATCCCTGACCTGGACGA 2111
Qy 315 uAlaAspMetIleLeuProIleValLeuGlnTyrLeuCysProValTyrThrIleSerPhe 335
Db 2110 GACGACATGATTTCTGCAATGAGTGGTGGAGTCTTCAACCTGACTTCGTGCTCTTCT 2051
Qy 335 eGlyLeuGlyAlaValSerAlaIaValMetSerSerAlaAspSerSerIleLeuSerAl 355
Db 2050 TGGATTGGGCGCTTTCGCGCGCGGTGATGTCTCGCGGACATCTCCGTTGCTTCGCG 1991
Qy 355 aSerSerMetPheAlaArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAsp 375

```
Db 1990 CGCCCTCCATGTCGCTCGGAACGTGTACAAATTGATTTCCCTCCGAGAGGCGCTCCGAGAT 1931
Oy 375 SGLIleValTTPValMetArgIleThrValIleuValPheGlyAlaSerAlaThAlaMe 395
Db 1930 GGAATTCATTGGGTGATGCGAGTCCGATCATTTGTGGGCAATCCGGTAACATCAT 1871
Oy 395 LAlaIleuLeuThrIlysrIhValTyrGlyLeuTPTr 407
Db 1870 GGCCCTCACCAATCCCTCCATCTACGTTGTGTG-GGTAGAGAAGATCATTTATTCACC 1812
Oy 408 -----LeuSe 409
Db 1811 CTCCTCCATCTATTTGCTAAGTCCCTGCTCTTCTTATCTTTCCGAGAGGCTCCATG 1752
Oy 409 rSerAspIleuValTyrIleIleIlePheProGlnIleuLeuCysValIleu---PheIlely 428
Db 1751 CTCGATCTGGTCTAGTCATCTCTTCCCGAGCTACTGATGTGTGCTCACTTCAAGAA 1692
Oy 428 sGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArgIleTh 448
Db 1691 GCACGCAACACAGTACGCGAGCGCTGTCGACATACATTGGCCCTGGCCATCCGACTGC 1632
Oy 448 rGlyGlyIleuProTyrIleuTyrLeuGlnProIleuIlePheTyrProGlyTyrTyrSerAs 468
Db 1631 GGGCGGTAGGCGCATCTTGGGAGCTGCTCCATTCATGATCAAGTATCCCGCTACAGAGGA 1572
Oy 468 pLysAsnGlyIleTyrAsnGlnArgPheProPheIlysrThrLeuSerMetValThrSerPh 488
Db 1571 GACCAAG-----GAGCAGATGTTCCCTTCCGACATAGGCGCATCTCTCAAGCT 1521
Oy 488 ePheThrAsnIleCysValSerTyrLeuAlaIys----- 499
Db 1520 GGTACGCTCATCTGCTCTCTGCTGGAGCTAAGTAAAGCTATGTAATTACAT 1461
Oy 499 ----- 499
Db 1460 AGTACTACATATTGATTATATTGTAATTAAATTAAATGTAACCTTCTCTTAC 1401
Oy 500 -----TyrLeuPheGlnSerGlyThrIleuProProlsLeuAspValPheAspAl 516
Db 1400 CAATTTCAGAAATGATGTTGATTCGGGCAAGTTGGCCGACACTACACTTCTCCGCTG 1341
Oy 516 aValVal 518
Db 1340 TGTGCTC 1334
RESULT 15
US-11-097-143-41594
; Sequence 41594, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
```

```
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41594
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-41594

Alignment Scores:
Pred. No.: 4,55e-135 Length: 1729
Score: 1405.00 Matches: 286
Percent Similarity: 69.1% Conservative: 71
Best Local Similarity: 55.3% Mismatches: 105
Query Match: 46.9% Indels: 55
DB: 13 Gaps: 8

US-10-724-806-4 (1-580) x US-11-097-143-41594 (1-1729)
Oy 4 HisValGlyGlyLeuValAlaIleIlePheTyrIleuIlePheLeuValGlyIle 23
Db 7 AATATCGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 66
Oy 24 TTPAlaIATrPLysThrLysAsnSerGlyAsnProGlnIlysrGlyAlaIleIle 43
Db 67 TGGCGCGGTCCGAG--AAGCAGTCCGCAATGATTCGAG--GAGGAGGTCATG 117
Oy 44 ValGlyGlyAspIleGlyLeuIleuValGlyGlyPheThrMetThrAlaThrVal 63
Db 118 CTGGCGGACGCTCCATGCGCTCTTCTGCGCATCTTACCATACGACCTGGGTG 177
Oy 64 GlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAla 83
Db 178 GGTGCGGCTCATATCAACGCGAGCGAGCTATATAC-----ACATCGGCTGCTG 231
Oy 84 TTPAlaHisAlaProIleGlyTyrSerLeuSerIleuIleuGlyGlyLeuPhePheAla 103
Db 232 TGTGTCCAGGCTCCATTGGATACCGCTTAAGCTTGATTTGGGCACTTCTTGGCC 291
Oy 104 LysProMetArgSerLysGlyTyrValThrMetLeuAspProPheGlnIleTyrGly 123
Db 292 AATCCATGCGAGAGGATTAACATCAACATGATGATCCGTTGACAGGATTCCTTGGT 351
Oy 124 LysArgMetGlyGlyLeuIleuPheIleProAlaIleuMetGlyGlyMetPheTrpAlaAla 143
Db 352 GAGCGGATGGAGAGATTCCTCTCTGCGCGCTATAGCGGTGAGGTCTTTGGGAGGCC 411
Oy 144 AlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAspValAsnIle 163
Db 412 GGCATCTGCTGCTCTCTCTGCAATCCCATCTTCAACAACCTTGGGTGAGCTGATCC 531
Oy 164 SerValIleValSerAlaIleuIleAlaIleuTyrThrLeuValGlyIleuTyrSer 183
Db 472 TGGTGATCTCTCTCTCTGCAATCCCATCTTCAACAACCTTGGGTGAGCTGATCC 531
Oy 184 ValAlaTyrThrAspValAlaGlnIleuPheCysIlePheIleGlyLeuTrpIleSerVal 203
Db 532 GTGGCGTATAGGAGATGATCATCATGTTGTCATCTTCACTGCTGAGTGCAT 591
Oy 204 ProPheAlaIleuSerHisProAlaValThrAspIleGlyPheThrAlaValHisAlaIys 223
Db 592 CCTTTCGCTGAGC----- 606
Oy 224 TyrGln-SerProTyrPleuGlyThrIleGlySerValGluValTyrThrTrpLeuAspAs 243
Db 607 ---AAGGAGCAGTGGGAGCGCTGAGTG-----ACCTGG----- 637
Oy 243 nPheLeuLeuIleuMetLeuGlyIleProTyrGlnAlaTyrPheGlnArgValIleuSe 263
Db 637 ----- 637
```

Oy	263	rSerserSerIaThrTyralaGlnValIleuSerPheunlaaIaPhegiyCysIleuVa	283
Db	638	-AGCAAAACGGCAGAGAGCGCCACCTCTCTCTATGTTGACGCGCCGATCATTTT	696
Oy	283	lMeaIaLeuProaIaIeCysIleGlyaIaIeGlyaIaSerThraSptIpaSgInTh	303
Db	697	GATGGCCATTCGCCCGGTGCTCATCGAGCATTTCCAAAGGCTACACTTGGAACGAC	756
Oy	303	rAlaTyrcIyTyrrProaSprProIySrrThlySgIuGlnuIaAspMetIleuProIleVa	323
Db	757	AGATTACAAAGGACCTTATCCCTGACCGTGGACAGACAGCAAGATTCGCCCATGT	816
Oy	323	lLeuGlnTyrrIleuCySProvaIlyrIleSerPhePhegiyLeuGlyaIaIaSerIaIaI	343
Db	817	GCTGCAGACTCTACGCGCTGACTTCGTCTCTCTTTGGATTGGCGCTTTCCGCGCC	876
Oy	343	aValMeSerSerAlaAspSerSerIleLeuSerAlaSerMetPhealAargAsnI	363
Db	877	CGTAGTCCTCCGCGCAGCTCTCGGTGCTCTCCGCGCCCTTCATGTTCCGTCCGAAGT	936
Oy	363	eTyrgIleuSerPhealargIlnaSnAlaSerAspIySgluIleValITrpValMetArgI	383
Db	937	GTACAAATGTATTTTCCGTGAGAAAGCCTCCGAGATGGAATCATTTGGGATGCGAGT	996
Oy	383	eThrrValLeuValPhegiyAlaSerAlaIleThAlaMerIaLeuIleuThrIySrrValTy	403
Db	997	CGCCATCTTTGTGGGGCATCTCTGGCTACATCATATGCGCTTCACCATTCCTTCATTA	1056
Oy	403	rGlyLeuTrpTyrrIleuSerSerAspLeuValTyrrIleIleIlePheProGlnLeuIeCy	423
Db	1057	CGATTGTGGTCCAGCTGCTCGGATCTGTCTACGTCATTTCTGCCGAGCTACGAT	1116
Oy	423	sValLeu---PheIleIySgIyThraSrrThTyrrcIyAlaValaIaGlyTyrrIlePheG	442
Db	1117	GGGTGTGACTTCAGAAAGACACTGCACACGACGAGCGCTGTGGATACATTGTGTC	1176
Oy	442	yLeuPheIleuArgIleThrGlyGlyGlnProTyrrIleuTyrrIleuGlnProIleIlePheTy	462
Db	1177	CTTGCCCATCCGATCGTGTGGGCGGTGAGGCCATCTTGGAGCTGCTCATTTGATTAAGTA	1236
Oy	462	rProGlyTyrrTyrrSerAspIySaSngIyIleTyraSngIlnArgPheProPheIyThLe	482
Db	1237	TCCCGGCTACGACGAGACCAAG-----GAGCAGATGTTCCCTTCGCGACCAT	1287
Oy	482	uSerMetValThrSerPheIleThraSrrIleCysValSerTyrrIleuAlaIyTyrrLeuPh	502
Db	1288	GGCCATGCTGTCAAGCTTCGTACGCTCATCTCGGTCTCTCCGTGAGCATAAAGATGATTT	1347
Oy	502	eGlnSerGlyThrIleuProProIySleuAspValPheAspAlaIleValI	518
Db	1348	TGAGTCCGGCAAGTTGCCGCCACGCTACGACTTCCGCTGTGGTGC	1396

Search completed: July 10, 2006, 14:47:49
Job time : 1929 secs

This Page Blank (uspio)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: July 10, 2006, 14:48:06 ; Search time 404 Seconds

(without alignments)
2799.791 Million cell updates/sec

Title: US-10-724-806-4

Sequence: 1 MFHVEGVAILFLIFL.....EALLDVSSPEGSGETDNIQ 580

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10724806/runatc.10072006.140301.7914/app_query.fasta.1
-DB=PubMed_Applcations_NA_New -QMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=bl0sum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs807
-USER=US10724806@CCN1.1137@runatc.10072006.140301.7914 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA_New:
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/BCT_NEW_PUB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	296	9.9	2449	US-11-266-748A-28918	Sequence 28918, A
2	296	9.9	2449	US-11-266-748A-28918	Sequence 28918, A
3	282.5	9.4	2273	US-11-266-748A-29027	Sequence 29027, A
4	257	8.6	20817	US-11-021-837-16	Sequence 16, Appl
5	255	8.2	1545	US-10-471-571A-1417	Sequence 1417, Ap
6	245.5	8.2	3227	US-11-266-748A-56747	Sequence 56747, A
7	235	7.9	95963	US-11-021-837-39	Sequence 39, Appl
8	206.5	6.9	2211	US-11-217-529-1082	Sequence 1082, Ap

9	205.5	6.9	2208	US-11-217-529-76899	Sequence 76899, A
10	197.5	6.6	1539	US-10-471-571A-1991	Sequence 1991, Ap
11	191	6.4	2984	US-11-266-748A-26072	Sequence 26072, A
12	184	6.1	38840	US-11-021-837-27	Sequence 27, Appl
13	169	5.6	3844	US-11-266-748A-27661	Sequence 27661, A
14	168	5.6	1578	US-11-266-748A-187885	Sequence 187885, A
15	143.5	4.8	936	US-11-266-748A-21747	Sequence 21747, A
16	143.5	4.8	936	US-11-266-748A-359132	Sequence 359132, A
17	143	4.8	933	US-11-266-748A-386971	Sequence 386971, A
18	143	4.8	933	US-11-266-748A-442511	Sequence 442511, A
19	143	4.8	933	US-11-266-748A-442511	Sequence 442511, A
20	142.5	4.8	2171	US-11-218-305-9650	Sequence 9650, Ap
21	137	4.6	1434	US-10-471-571A-3889	Sequence 3889, Ap
22	136	4.5	2513	US-10-449-902-20129	Sequence 20129, A
23	133.5	4.5	1966	US-11-266-748A-27363	Sequence 27363, A
24	131	4.4	59143	US-11-021-837-34	Sequence 34, Appl
25	129.5	4.3	2429	US-10-449-902-23815	Sequence 23815, A
26	127	4.2	1000	US-11-266-748A-291922	Sequence 291922, A
27	127	4.2	1000	US-11-266-748A-34351	Sequence 34351, A
28	127	4.2	1000	US-11-266-748A-403483	Sequence 403483, A
29	127	4.2	1000	US-11-266-748A-474529	Sequence 474529, A
30	124	4.1	1209	US-10-471-571A-2713	Sequence 2713, Ap
31	123	4.1	1730	US-11-174-3078-5197	Sequence 5197, Ap
32	122.5	4.1	6570	US-11-021-837-10	Sequence 10, Appl
33	121	4.0	1928	US-10-449-902-20061	Sequence 20061, A
34	121	4.0	3070	US-11-218-305-1841	Sequence 1841, Ap
35	120	4.0	1377	US-10-471-571A-471	Sequence 471, Ap
36	119.5	4.0	2614	US-11-216-545-8626	Sequence 8626, Ap
37	119	4.0	2186	US-11-266-748A-33772	Sequence 32772, A
38	119	4.0	2237	US-11-266-748A-23898	Sequence 23898, A
39	118.5	4.0	1000	US-11-266-748A-288126	Sequence 288126, A
40	118.5	4.0	1000	US-10-449-902-239555	Sequence 239555, A
41	117.5	3.9	2216	US-10-449-902-24763	Sequence 24763, A
42	117	3.9	1592	US-11-216-545-3055	Sequence 3055, Ap
43	116.5	3.9	1941	US-11-174-3078-1767	Sequence 1767, Ap
44	116.5	3.9	100445	US-11-021-837-40	Sequence 40, Appl
45	116	3.9	1470	US-11-217-529-396	Sequence 396, Appl

ALIGNMENTS

RESULT 1
US-11-266-748A-28918
; Sequence 28918, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkins, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662, 276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700, 293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28918


```

RESULT 2
US-11-266-748A-56547
/ Sequence 56547, Application US/11266748A
/ Publication No. US2006013463A1
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ PRIOR FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO: 56547
/ LENGTH: 2449
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-56547

Alignment Scores:
Pred. No.: 3,45e-22 Length: 2449
Score: 296.00 Matches: 148
Percent Similarity: 38.1% Conservative: 99
Best Local Similarity: 22.8% Mismatches: 222
Query Match: 9.9% Indels: 180
DB: Gaps: 29

US-10-724-806-4 (1-580) x US-11-266-748A-56547 (1-2449)
QY 11 ILeileuPheTyrrLeuLeuilePheLeuValGlyIleThrAlaAlaTrpLysThrLys 30
DB 104 ATCGTATTACTCTCTGTAAGATGCGCGTGGACGTGGGCTATGTTTCCAC--- 160
QY 31 AAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyGlyArgAspIleGly 50
DB 161 AATCGGGAGACT-----GTTGAGGCTTCTCTCGGCAGGCGCAAGTATGATG 208
QY 51 LeuLeuValGlyGlyPheThrMetThrAlaThrTrpValGlyGlyLysTrpIleAsnGly 70
DB 209 TGGTGGCCGATGGACCTCCCTCTTGGCTAGTACATGGAGAGTGCACCTTGTGGGG 268
QY 71 ThrAlaGluAlaValTyrrGlyProGlyGlyLeuAlaTrpAlaHisAlaProIleGly 90
DB 269 CTGGCC-----GGGACTGGGGCAGCTTCACGCAATCCGCAATGGAGGC 310
QY 91 TyrSer-----LeuSerLeuIleLeuGlyGlyLeuPhePheAlaLysPro 105
DB 311 TTGGAATGGAATGCCCTGTTGTTGGTGGTGTGCTGGGCTGGCTGTTTTC---CCC 364
QY 106 MetArgSerLys---GlyTyrrValThrMetLeuAspProPheLysGlnIleTyrrGlyLys 124
DB 365 ATCTATTATTAAGGCTGGGGGTGGTGAACAATGCCAGAGTACTGAGG-----AAG 412
QY 125 ArgMetGlyGly-----LeuLeuPheIle 132

```

```

DB 413 CGGTTTGAGGCGCAGGATCCAGGCTTACTTCCCTTCTGTCCTGCTGCTCTACATT 472
QY 133 ProAlaLeuMetCylGluMetPheThrAlaAlaIlePheSerAlaLeuGlyAlaThr 152
DB 473 TTCACCAAGATCTCGGACACATCTCTCGGGGGCCATATC----- 514
QY 153 IleSerValIleIleAspValAspValaIleIleSerValIleValSerAlaLeuIleAla 172
DB 515 ATCATCTGGCCTTAGGCTGATCTGATTATAGCATCTTCTCTTATTGGCAATCACT 574
QY 173 IleTyrrThrLeuValGlyGlyLeuTyrrSerValAlaTyrrThrAspValValGlnLeu 192
DB 575 GCCCTTACACAAATTACAGGGGGGCGTGGCGGTATTATACAGGACACCTTGACAGC 634
QY 193 PheCysIlePheIleGlyLeuTrpIleSerValProPheAlaLeuSerHisProAlaVal 212
DB 635 GTGATCATGCTGGTGGGCTTTAATCTGACTGGGTGCTTTTACAGAACTGGGA--- 691
QY 213 ThrAspIleGlyPheThrAlaValHisAlaLysTyrrGlnSerProThrLeuGlyThrIle 232
DB 692 -----GGCTATGACGCTTCAATGAAAAGTACATGAAGCC---ATTCACACATA 739
QY 233 GluSer-----ValGluValTyrrThr---TrpLeuAspAsnPhe 244
DB 740 GTGCTGATGGAACACACACCTTTCAGAAAATGCTACATCCACAGGCGGCACTCTTC 799
QY 245 LeuLeuLeu-----MetLeuGlyGlyIleProTrp----- 254
DB 800 CACATCTTCCAGATCCCTCAAGGAGACCTCCATGCGCTGGTTCATCTTTGGGATG 859
QY 255 -----GlnAlaTyrrPheGlnArgValLeuSer 263
DB 860 TCCATCTTACCTTGGTGTGATCTGTGTCACAGATCAGCTCATCTGACGCGCTCTCA 919
QY 264 SerSerSerAlaThrTyrrAlaGln-----ValLeuSerPheLeuAlaAlaPhe 279
DB 920 GCCAAGATATGCTACACGTGAAGGGGTGGCTGATCTGTGGGTATCTAAAGCTGATG 979
QY 280 GlyCysLeuValMetAlaLeuProAlaIle----- 289
DB 980 CCCATGTTCATGATGTGATGCGCAGGAATGATCAGCCGCAATTCTGTACACAGAAAATT 1039
QY 290 -----CysIleGlyAlaIleGlyAlaSerThrAspTrpAsnGln 302
DB 1040 GCCTGTGCTGCCCTTCAGAAATGTGAATAATTTGGGTATCCAAAGTGGCTTACCAAC 1099
QY 303 ThrAlaTyrrGlyTyrrProAspProLysThrLysGluGluAlaAspMetIleLeuProIle 322
DB 1100 ATGGCCTAT-----CCAAAC 1114
QY 323 ValLeuGlnTyrrLeuCysProValTyrrIleSerPhePheGlyLeuGlyAlaValSerAla 342
DB 1115 TAGTGTGGAGCTCATGCGCCCAATGAGCTGCAGGCGCTGATGTATCATGCTGCGGCC 1174
QY 343 AlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsn 362
DB 1175 TCCCTATATAGCTCCCTGACCTTCATCTTCAACAGGCGCAGCACTCTTCCACCAATGAC 1234
QY 363 IleTyrrGlnLeuSerPheAlaGlnAlaAspAspLysGluIleValTyrrValMetArg 382
DB 1235 ATCTAC---GCCAAGTCCGCAAGAGAGCATCTGAGAAAGAGCTCATGATGGCGGAGG 1291
QY 383 IleThrValLeuVal---PheGlyAlaSerAlaThrAlaMetAlaLeuThrLysThr 401
DB 1292 TTGTTTATCTGGTGTGATGGCATCAGCATGCGCTGGGTGCCATTGGCAGTCAACA 1351
QY 402 ValTyrrGly-----LeuTrpTyrrLeuSerSerAspLeuValTyrrIle-----IleIle 417
DB 1352 CAAAGTGGCAACTTTCGATTATCATCATGCTCATCAGCATGATCTGGGACCAACCATT 1411
QY 418 PheProGlnLeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrrGlyAlaValAla 437

```

Db	1412	GC	GC	CTGCTTCC	TCGCTTGCTATTTT	TCGAAAGAGAGCATGAGCCAGAGCC	TTTGG	1471																			
Qy	438	G	L	Y	T	R	I	L	E	P	H	E	G	L	E	P	H	E	A	R	G	I	E	-----	447		
Db	1472	G	G	A	C	T	G	A	T	C	T	T	A	G	A	C	T	T	G	A	T	T	---	1531			
Qy	448	T	H	R	G	L	Y	-----	G	L	Y	-----	G	L	Y	-----	G	L	Y	-----	G	L	Y	-----	454		
Db	1532	A	C	C	G	G	A	G	A	C	T	G	A	G	C	C	A	G	A	C	A	T	T	---	1591		
Qy	455	T	Y	L	E	U	G	I	N	P	R	O	L	E	U	I	L	E	P	H	E	T	Y	-----	474		
Db	1592	T	A	C	T	T	G	C	C	A	T	T	A	C	T	T	T	-----	1612								
Qy	475	G	L	N	A	R	P	H	E	P	R	O	P	H	E	L	S	E	R	V	E	R	V	-----	494		
Db	1613	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	494			
Qy	495	S	E	R	T	Y	R	E	U	A	L	S	T	Y	R	E	U	P	H	E	G	L	S	E	R	514	
Db	1643	T	C	C	C	T	C	T	C	A	C	C	A	A	C	C	C	A	T	-----	-----	-----	-----	-----	1681		
Qy	515	A	S	P	L	A	V	A	L	-----	V	A	L	A	A	R	G	I	S	E	R	G	L	S	E	532	
Db	1682	C	G	T	C	T	G	T	T	G	A	C	T	C	T	C	G	A	C	A	A	G	A	G	A	1735	
Qy	533	A	R	G	A	N	G	L	A	N	I	L	E	U	S	A	N	G	L	U	E	N	A	L	P	-----	552
Db	1736	G	A	M	A	G	A	G	A	A	C	A	T	C	A	A	G	A	G	A	G	A	G	-----	1777		
Qy	553	L	E	U	S	E	R	T	H	R	P	H	E	T	A	S	I	N	L	E	G	L	-----	561			
Db	1778	A	T	A	G	A	A	C	A	C	A	G	T	T	C	T	C	G	A	G	A	A	-----	1804			

Score:	282.50	Matches:	143
Percent Similarity:	36.2%	Conservative:	92
Best Local Similarity:	22.0%	Mismatches:	240
Query Match:	9.4%	Indels:	175
DB:	8	Gaps:	23
US-10-724-806-4 (1-580) x US-11-266-748A-29027 (1-2273)			
Qy	8 LeuValAlAllelleuPheTyLeuLeuIlePheLeuValGlyIleTPraAlaIleTyr	27	
Db	96 ATCTGATGATTCGTCGATATTTCTGCTGGTCATGGCCGTGGCTTGGTGCATGTC	155	
Qy	28 LysThrLysAsnSerGlyAsnProGluIlyArgSerGluAlaIleIleValGlyArg	47	
Db	156 AGAAC--AACAGAGGCACT-----GTGGCGGCTACTTCCTGACAGAGC	200	
Qy	48 AsprIleGlyLeuLeuValGlyIlePheThrMetThrAlaThrTPraValGlyGlyIlyTyr	67	
Db	201 AGCATGCTGTGGTGGCCGGTGGGCTCTCTCTCTGGCCACCAACATGGAGTGGCCAC	266	
Qy	68 IleAsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaTPraIleVala	87	
Db	261 TTGTGGGCGCTGGCAGAGCACT-----GGCGGTGAATGGCTTGGCTGGTGCAGATTC	314	
Qy	88 ProIleGlyTyrSerLeuSerLeuIleLeuGlyIlePhePheIleAllySPrometArg	107	
Db	315 GAGTGAAATGCGCTCTTCGTGGTGGTGTACTAGGCGCTGGCTGTGGACCGGTACTCG	374	
Qy	108 SerLysGlyTyrValThrMetLeuAspProPheLysGlnIleTyrGlyLysArgMetGly	127	
Db	375 ACAGCGGGGTATCATCGATG-----CCAGTGATCGCGCAAGCGCTTGGC	422	
Qy	128 Gly-----LeuLeuPheIleProAlaLeu	135	

```

RESULT 3
US-11-266-748A-29027
Sequence 29027, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Hartlin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptome Microarray Technology and
TITLE OF INVENTION: Methods of Using the Same
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11266, 748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 48396
SOFTWARE: PatentIn version 3.3
SEQ ID NO 29027
LENGTH: 2273
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-29027

```

[illegible]

```

Db      987 AGCCCATGTTCTCATGTCATGCGAGCATGATCAGCGCGCATCTGTATCCAGACGAG 1046
Qy      280 ---GlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThr 298
Db      1047 GTGGCGGTGGGTG---CCTGAGGTGTGAGCGCGCTGTGCGGACGAGGTG 1097
Qy      299 AspTPasAsnIleThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMet 318
Db      1098 GCGTGTCCACATCGCCTAC----- 1118
Qy      339 IleLeuProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGly 338
Db      1119 ---CCGCGGTCTGCTGAGATCATGCGCAACGCGTGTGCGGAGTATCATGCGG 1172
Qy      339 AlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerMet 358
Db      1173 GTATGCTGCGCGCTCATGTCCTGCTGCGCTTCATCAACAGCAGCAGCAGCTC 1232
Qy      359 PheAlaArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleVal 378
Db      1233 TTCACCATGACATCTAC---ACGCGCTGCGGCGCACGCGCGGCGACCGCGAGCTGCTG 1289
Qy      379 TrpValMetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeu 398
Db      1290 CTGGTGGAGCGGCTGTGGGTGTTCATGCTGATGTCGTGGTGGCTTCCGCTTCCGCTG 1349
Qy      399 ThrLysThrValTyrGly-----LeuTyrTyrLeuSerSerAspLeuValTyrIle 415
Db      1350 GTGACGCGCGCAGCAGCGCGGCGAGCTTCGATTACATCAGGACATCTCTAGCTACTG 1409
Qy      416 Ile-----IlePheProGlnLeuLeuCysValLeuPheIleLysGlyThrAsnThrTyr 433
Db      1410 GCACCGCCCGTGTCCCGCTTCGTCGTCGCGCTTCGTCGCGCGCGCTTATGACGAG 1469
Qy      434 GlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArgIleThrGly----- 449
Db      1470 GCGCGCTTCTGGGGACTCATCGGGGCGCTGATGAGCGCTGCGACGCTGATCCCGAG 1529
Qy      450 -----Gly 450
Db      1530 TTCTCTTGGCTCGGCGAGCTGTGTCAGCCCTCGCGCTTCCCTGTCGCGC 1589
Qy      451 GluProTyrLeuTyrLeuGlnProLeuIlePheTyrProGlyTyrTyrSerAspLysAsn 470
Db      1590 GTGCACTACTCTACTCTCCCATTTGCTGCTTCTC----- 1625
Qy      471 GlyIleTyrAsnGlnArgPheProPheLysThrLeuSerMetValThrSerPhePheThr 490
Db      1626 -----TGCTGTGCGCTCTCTCAACCTTCACGCTC 1652
Qy      491 AsnIleCysValSer-----TyrLeuAlaLysTyrLeuPheGluSerGly 505
Db      1653 TCCCTGTGACCGCGCCATCCCGAGAAACACCTTCACCGCTGTGCTTCACTCTCCG 1712
Qy      506 ThrLeuProProLysLeuAspValPheAspAlaValAlaLysIleSerGluGluAsn 525
Db      1713 CATAGAGAGAGAGAGAGAGAGAGCTGATGCT-----GATGAGCAG 1754
Qy      526 MetAspLysThrIleLeuValArgAsn-----GluAsnIleLysLeuAsnGlu 541
Db      1755 CAAGGCTCTCACTCTCTGACAGATGGGTGCCAGAGAGTGCATGAGATGATGAG 1814
Qy      542 LeuAlaProValLysProArgGlnSerLeu 551
Db      1815 -----CCCCAGGCGCGGACCAAGCCTC 1838

```

```

; APPLICANT: Manning, Paul
; APPLICANT: McLaughlin, Robert
; APPLICANT: McCormack, Kathleen
; TITLE OF INVENTION: Staphylococcus Saprophyticus Nucleic Acids and Polypeptides
; FILE REFERENCE: 101324-US
; CURRENT APPLICATION NUMBER: US/11/021,837
; PRIOR FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/533534
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/600660
; PRIOR FILING DATE: 2003-08-31
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 16
; LENGTH: 20817
; TYPE: DNA
; ORGANISM: Staphylococcus saprophyticus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15033)..(15033)
; OTHER INFORMATION: n is a, c, g, or t
US-11-021-837-16

Alignment Scores:
Pred. No.: 1.8e-16 Length: 20817
Score: 257.00 Matches: 120
Percent Similarity: 42.4% Conservative: 109
Best Local Similarity: 22.2% Mismatches: 215
Query Match: 8.6% Indels: 96
DB: Gaps: 23

US-10-724-806-4 (1-580) x US-11-021-837-16 (1-20817)
Qy      9 ValAlaIleIleLeuPheTyrLeuLeuIlePheLeuValGlyTyrTrpAlaIleTrpLys 28
Db      18837 ATCAAGCTTGGTGGTATTTTGTGGTGTGCTAGTATGATGGACTATGCTT----- 18890
Qy      29 ThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyArgAsp 48
Db      18891 AAGCATCTACCGGTATACCTGAGTGA-----TATATGCTGGCGGCTAGAAAC 18938
Qy      49 IleGlyLeuLeuValGlyGlyPheThrMetThrAlaThrTyrValGlyGlyTyrIle 68
Db      18939 ATGGTCATATGATGATCATCATATTCGACGCGCATGACATGATGATGATGATGAT 18998
Qy      69 AsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaTrpAlaHisAlaPro 88
Db      18999 ATGGGATTACCTGTGATATATATATACCGGTTTATACGACGCTGG-----TTAGCG 19052
Qy      89 IleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeuPhePheAlaLysProMetArg--- 107
Db      19053 ATTTGATTACACTTGGCGCATATACATATATATTTGTTGTTGTTGTTGTTGTTGTTG 19112
Qy      108 -----SerLysGlyTyrValThrMetLeuAspProPheLysGlnIleTyrGly 123
Db      19113 TATACAGAAAGAGCGGAGATGCAATTAACATCACTGATTTCTTTAAATAAGATTAGAT 19172
Qy      124 LysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPheTrpAlaAla 143
Db      19173 GATTAATCTTAATTCGATTAATAATTTATTTACAGTGCATTAATGTTATTTTTCACACT 19232
Qy      144 AlaIlePheSerAlaLeu-----GlyAlaThrIleSerValIleIleAspValAsp 160
Db      19233 TATACCATTTCTGGATGTATGATCAGCGGTTGATTTTGAAGAAGTCGTTGGCGTTAAC 19292
Qy      161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
Db      19293 TATCACTATGATGATATCTTGTGCTGATGATGATGATGATGATGATGATGATGATGAT 19352
Qy      181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db      19353 TATTTAGCAGTATCATTAACAACATTTTTCAGAGGTGATGATGATGATGATGATGATG 19409

```

```

RESULT 4
US-11-021-837-16
; Sequence 16, Application US/11021837
; Publication No. US20060140972A1
; GENERAL INFORMATION:
; APPLICANT: Alm, Richard

```

```
OY 201 I l e s e r V a l p r o P h e a l a l e u s e r H i s p r o a l a v a l t h r a s p l l e g l y p h e t h r a l a v a l 220
Db 19410 G T T A G T A C C A - - - - - A T A 19424
OY 221 H i s a l a y t y r g l n s e r p r o t r i p l e u g l y t h r i l e - - - G l u s e r V a l g l u v a l y t y r - - - 238
Db 19425 G T T G C T A T G T T G C A A T T A A G C G A C T T G A T T A C C T T T T A C G C A A C A G C A A T T A A A A C G G 19484
OY 239 T h r i t r i p l e u a s p a n s p h e - - - - - L e u l e u M e t l e u g l y g l y l e p r o t r p 254
Db 19485 A C A A C T T A G A C T T G T T A A A G A A C A G C A T T A T A G A T T A T T C A T T C T T G G C T G G 19544
OY 255 - - - G l n a l a t y r p h e g l n a r g - - - - - V a l l e u s e r S e r S e r a l a t h r t y r a l 270
Db 19545 G G A T T A G G T T A C T T G T T G C A C C A C A T A T T A T T A C G T T T A T T A C A T T A A T C T G T C 19604
OY 271 G l n a l e u s e r p h e l e u a l a l a p h e g l y c y s l e u a l M e t a l e u p r o a l a l e c y s 290
Db 19605 A A A C A T T A C C A A C G C A A G A C G T T T G G T A T C G G T T G A T G C C A A T C A G C - - - - - T T A 19658
OY 291 I l e g l y a l a l e g l y a l a s e r t h r a s p t r p - - - - - A s n g l n t h r a l a t y r g l y 306
Db 19659 T T A G C G C A G T G G T G T G C G A T T G A C A G G A T T G C G T T A T T A T T A A A G C G T A C A G A T 19718
OY 307 T y r p r o a s p r o l y s e r t h r l y s e g l u a l a a s p m e t l e u e u p r o l l e v a l l e u g l n t y r 326
Db 19719 A T A G A A A A T C C T G A A A C A - - - - - C T A T T C A T T T T A T T A T G G C A C A A A T T 19760
OY 327 L e u c y s p r o v a l t y r l l e s e r p h e g l y l e u g l y a l a l a s e r a l a l a l a l a l e u s e r 346
Db 19761 T T A T T C C A T C C C T T A G T T G T G A T T C T T G C T T G C T A T T T T A T T T A G C G A T T A G A C T 19820
OY 347 S e r a l a a s p s e r S e r l l e u s e r a l a s e r m e t p h e a l a a r g a s n l l e y t r g l n e u 366
Db 19821 A C A A T T T C T T C A C A T T G T T A T T A A C G T C A A G T T C T T A A C A G A A C T T T A T T A A A T A 19880
OY 367 S e r p h e a r g g l n a s n a l a s e r - - - - - A s p l y s g l u l e v a l t r p a l a l e u m e t a r g l i e 383
Db 19881 A T T C G T G T G A G A A G C T G C A A G A A C A T G A A A A A G A A T T T G T T A C T A G C G A T T A 19940
OY 384 T h r a l l e u a l p h e g l y a l a s e r a l a t h r a l a m e t a l e u l e u t h r y s t h r a l y t y r 403
Db 19941 T C A G I C A T C A T C - - - - - G T A G C A T T G T - - - - - T C G A T T T G G 19973
OY 404 G l y l e u t r p t y r l e u s e r S e r a s p l e u a l t y r l l e i l e - - - - - 416
Db 19974 A T T G C T T G C A C C T A A C A T A C A T T T T G G C T C T A G T A A T A T G C A T G G C T G T T C 20033
OY 417 - - - - - l l e p h e p r o g l n e u l e u c y s v a l l e u p h e i l e y g l y t h r a n t h r t y r 433
Db 20034 G G T G C C G A T T G G A C C A C T T A C T T A T C C T T A T A T T G A A G G G T T A A G T A G A C A 20093
OY 434 G l y a l a l a l a g l y t y r l l e p h e g l y l e u p h e l e u a r g l l e t h r c l y g l y g l u p r o t y r 453
Db 20094 G G T G C T G A A G C G A T T A T C A G G T G C A T T G T A T G T A T C A T A T G - - - - - A T T 20144
OY 454 L e u t y r l e u g l n p r o l e u l l e p h e t y r p r o g l y t y r S e r a s p l y s a n g l y l l e t y r 473
Db 20145 G C A T T G T T A A C C A T T A - - - - - G G T G A T G A A T A T T C T T 20183
OY 474 A s n g l n a r g p h e p r o p h e l y s t h r l e u s e r M e t V a l t h r S e r p h e t h r a n l l e c y s 493
Db 20184 A A T - - - - - T T A T A T G A A A T T A A C T G A T T T T T A C T G A T T A A T C 20225
OY 494 V a l s e r t y r l e u a l a l y t y r l e u p h e g l u s e r g l y t h r l e u p r o p r o l y s e u a s p a l 513
Db 20226 G T T A C G T T T A G T A A G T A A T T C - - - - - A C G A A A A C C A C A A T T A T G A T T 20273
```

```
APPLICANT: CHIRON SPA
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE: P02692720
CURRENT APPLICATION NUMBER: US/10/471,571A
PRIOR FILING DATE: 2003-09-12
PRIOR APPLICATION NUMBER: GB-0107661.1
NUMBER OF SEQ ID NOS: 5642
SOFTWARE: Seqwinn9, version 1.03
SEQ ID NO 1417
LENGTH: 1545
TYPE: DNA
ORGANISM: Staphylococcus aureus
US-10-471-571A-1417

Alignment Scores:
Pred. No.: 5,34e-18 Length: 1545
Score: 255.00 Matches: 124
Percent Similarity: 40.1% Conservative: 89
Best Local Similarity: 23.4% Mismatches: 220
Query Match: 8.5% Indels: 98
DB: 6 Gaps: 24

US-10-724-806-4 (1-580) x US-10-471-571A-1417 (1-1545)
OY 9 V a l a l a l l e l e u p h e t y r l e u l e u l e p h e l e u a l g l y l e t r p a l a l a t r p l y s 28
Db 67 A T T A T A T T G C C G T T A C T T C T T G T T A C T A A T C G T A T T G G C T T T A C G G T - - - - - T A C 120
OY 29 T h r l y s a n s e r g l y a s p p r o g l u g l n a r g s e r g l u a l l e l e v a l g l y a r g a s p 48
Db 121 A A G C A A G A C A C T G T A A C C T T A A C G A G - - - - - T A C A T G T T A G G T G A C G A C T G T 168
OY 49 I l e g l y l e u l e u a l g l y l y p h e t h r m e t t h r a l a t h r p a l a l y g l y g l y t y r l l e 68
Db 169 A T T G A C C G T A T A T A C G A T T A C G A T T A C A C T G A C T T C A G A T A T A G A G A T G A T G A T T 228
OY 69 A s n g l y t h r a l a g l u a l a v a l t y r g l y p r o g l y c y s g l y l e u a l t r p a l a h i s a l a p r o 88
Db 229 A T G G G C T A C C T G T T C T G T A T A T A G C A C T G C T A T A C A G C A T A T G C - - - - - A T T A C A 282
OY 89 I l e g l y t y r s e r l e u s e r l e u l e u g l y l e u p h e p h e a l a l y s p r o m e t a r g s e r 108
Db 283 A T C G G T T T A C A T T A G T G C T T A T A T A A T - - - - - T A C T T G T T G T T G C T C T A G A C T T 336
OY 109 L y s g l y t y r - - - - - V a l t h r m e t l e u a s p p r o p h e l y s g l n l l e 121
Db 337 C G T G T T T A T A C C G A T T A G C T G A G A T G C A A T T A C A T T A C C A G A T T C T T T A A A A T G T 396
OY 122 T y r g l y l y s a r g m e t g l y l y l e u l e u p h e i l e p r o a l a l e u m e t g l y g l u m e t h e t r p 141
Db 397 T T A A A C G A T A A A A A T A A T A T G T T A A A G A T T A T T C T G A T G A T A T G T A T T C T T 456
OY 142 A l a l a l a l l e p h e s e r a l e u - - - - - G l y a l a t h r l l e s e r V a l l l e l e a s p 158
Db 457 A C A T T A T A T A C A C A T T C G G T T C G T A C T G G T G A A A C T A T T T G A A A G C C T T T G G A 516
OY 159 V a l a s p a l a s n l l e s e r V a l l l e v a l s e r a l a l e u l l e a l l e l e u t y r t h r l e u a l 178
Db 517 T T A G A T T A T A C A T T T G G T T A A T A T A T A T A T T A G T G C T T A C A T T T C T T A A T C T T C T T 576
OY 179 G l y g l y l e u t y r s e r V a l a l a t y r t h r a s p a l V a l g l n l e u p h e c y s l l e p h e i l l e g l y 198
Db 577 G G T G A T T A T T A G C T G T A C A A T T T C C A A A G G T G C A T T A T T A T T A T T A T T A T T A T T 636
OY 199 L e u t r i p l i e s e r V a l p r o p h e a l a l e u s e r H i s p r o a l a v a l t h r a s p l l e g l y p h e t h r 218
Db 637 A T G - - - - - G T T A G T G C C C A T T G T T - - - - - G C A T A G A T A A T T A - - - - - 672
OY 219 A l a V a l H i s a l a l a t y r g l n s e r p r o t r i p l e u g l y t h r l l e g l u s e r V a l g l u v a l y t y r 238
Db 673 - - - - - A A C G C T G G - - - - - G G A A C G T T C A G A T A C A C A C A T A T G 708
```



```

QY 239 -----ThrrpLeuaspasnPhenLeuLeuMetLeuGlyGlyIle----- 252
DB 709 AAACCCACAAATTTAATTATTTAAAGGTTATCATCTTATAGATTATCTCTATTT 768
QY 253 ProTrp--GlnAlaTyrrheGlnArg-----ValLeuSerSerSerSerAlaThr 268
DB 769 TCATGGGGATTAGTATTTGGTCAACCTCATATCATCTTAAAGTTATGTTATTTAA 828
QY 269 TyrrAlaGlnValLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAla----- 285
DB 829 TCACACACATGCTACCTTAAGCTAGACCTTTAGGATTTAGCTGATGCTGTTTAA 888
QY 286 LeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrpAsnGlnThrAlaTy 305
DB 889 TTAGGGCGCTGGCTGTTGTTTAAACAGTATTTGCATTC-----GTACCTGCTTAT 939
QY 306 -----GlyTyrrProaspProlysthrLysGluGluAlaAspMetIleLeuProIleVal 323
DB 940 CATATTAAACTAGAAAGATCTCGAGACA-----TTATTCAATCGTGATG 981
QY 324 LeuGlnTyrrLeuCysProValTyrrIleSerPhePheGlyLeuGlyAlaValSerAlaAla 343
DB 982 AGTCAGATCTCTTCATCTCTTGTAGGTGCTTTTACTTCTGCTCGATTCAGCTGCA 1041
QY 344 ValMetSerSerAlaaspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIle 363
DB 1042 ATTATGAGCAGCATTTCTTCACAAATTACTGTAACTAGTTCACATTAACGAGACTTT 1101
QY 364 TyrrGlnLeu-----SerPheArgGlnAsnAlaSerAspLysGluIleValTyrrVal 380
DB 1102 TATAAATTAATTCTGTGTGTAAGAAAGAAAGCTAAACCCACCAAAAGAAATTTTATGAT 1161
QY 381 MetArgIleThrrValLeuValPheGlyAlaSerAlaThrrAlaMetAlaLeuThrLys 400
DB 1162 GGAAGATTATCTGTATTTAGTGTAGCAATTGTTGCCATCGCATGCA----- 1209
QY 401 ThrrValTyrrGlyLeuThrTyrrLeuSerSerAspLeuValTyrrIleIle----- 416
DB 1210 -----TGGAAATCCAAACGACACACATTTCTAACTTAAGTAAGTAACGCTTGG 1254
QY 417 -----IlePheProGlnLeuLeuCysValLeuPheIleLysGlyThr 430
DB 1255 GCCGGATTGGTCATCGTTACAGTCCACTTGTCTATTGACCTTAACTGAGAAAGCTTGG 1314
QY 431 AsnThrTyrrGlyAlaValAlaGlyTyrrIlePheGlyLeuPheLeuArgIleThrGlyGly 450
DB 1315 ACACGTGCGGTCGTGTAGATGGAATGTTTCAGGTGCTTAACTGTTATCGTTTGG-- 1371
QY 451 GluProTyrrLeuTyrrLeuGlnProLeuIlePheTyrrProGlyTyrrTyrrSerAspLysAsn 470
DB 1372 -----ATTGCATGCAATTAAACCATTTG-----GCACATTCACAGAAATATTC 1413
QY 471 GlyIleTyrrAsnGlnArgPheProPheLysThrLeuSerMetValThrSerPhePheThr 490
DB 1414 GCCTTATAT-----GAAATTATCTCGATTTATTGTA 1446
QY 491 AsnIleCysValSerTyrrLeuAlaLysTyrrLeu 501
DB 1447 AGGTATATCGTTACATATGTGTAAAGTAAACTT 1479

```

```

; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56747
; LENGTH: 3227
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-11-266-748A-56747

Alignment Scores:
Pred. No.: 1,84e-16 Length: 3227
Score: 245.50 Matches: 125
Percent Similarity: 36.2% Conservative: 79
Best Local Similarity: 22.2% Mismatches: 162
Query Match: 8.2% Indels: 197
Gaps: 27
DB:

US-10-724-806-4 (1-580) x US-11-266-748A-56747 (1-3227)
QY 4 HisValGlnGlyLeuValAlaIleIleLeuPheTyrrLeuLeuIlePheLeuValGlyIle 23
DB 105 CACGCTACGACATACAGCGCTGTGTCATCTTGTCTGCTCATTCCTGCGGATC 164
QY 24 TrpAlaAlaTyrrLysThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIle 43
DB 165 TGGTGTCTCATCTGCGCA-----AGTGGAGGACCAT 197
QY 44 ValGly-----GlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThr 59
DB 198 GCGCGTATTTCTGCGCGGAGGTCCTACAGCTGTGCGCAATTCGACATCTCGATG 257
QY 60 AlaThrTrpValGlyGlyGlyTyrrIleAsnGlyThrAlaGluAlaValTyrrGlyProGly 79
DB 258 TCACGAAATGTGGGAGGTGCTGTTTCATCGGCTGGCT-----GGACAGGG 305
QY 80 Cys-----GlyLeuAlaTrp----- 84
DB 306 GCTGCCGAGGCTTGGCTGAGTGTGCTGAGATGACATGAGAAATCAAGCTGTGA 365
QY 85 -----AlaHisAlaProIleGlyTyrrSer----- 92
DB 366 GGAAGACAGAGGATTCATCCAGTCAAGTCACGAGAGACTGGGGTCAAGTCCAGCAACC 425
QY 93 ---LeuSerLeuIleLeuGlyLeuPhePheAlaLysProMetArgSerLysGlyTyrr 111
DB 426 TGGCTGCTCTGCGCTTGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 482
QY 112 ValThrMetLeuAspProPheLysGlnIleTyrrGlyLysArgMetGlyGly----- 128
DB 483 GTACATATG-----CCGCGATTCGAAAGACGATTTGGGGCCAGAGATC 530
QY 129 -----LeuLeuPheIleProAlaLeuMetGlyGluMet 139
DB 531 CAGGTATACATGCTGTCTGCTCTCATCTCTACATCTTCAACCAAGATTCGACTGAC 590
QY 140 PheTrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspVal 159
DB 591 ATCTTCTCGAGGCTCTTC-----ATCCAGATGCAATTTGGCTGG 632

```

QY 160 AspvAlaenIIeSeValIIeValSerAlaIeulIeAlaIleLeuTyThrIeuValGly 179
Db 633 AACCTGACCTCTCCACAGAGCTGCTGAGTACGCGCTACACCATTCAGGT 692
QY 180 GlyLeuTyThrSerValAlaIeTyThrAspValValGln 191
Db 693 GGCCTCATGGCGGTATCTACACAGATGCTGACAGCGTGATCATGTAGGGGAGCC 752
QY 192 LeuPheCysIlePheIleGlyLeu 200
Db 753 CTGGTCCATGTTCTGCGCTTTAGGACGCGGCTGTACCCAGGCTGAGAGAGCGG 812
QY 201 1leSeValIProPheAlaIeSerHis 209
Db 813 TACAGGAGGCCATCCCTAATGTCACAGTCCCAACACCCTGTACCTCCAGCGCC 872
QY 210 ProAlaValThrAspIleGlyPheThrAlaValHis 221
Db 873 GATGCTTCCACATTCTTGCGGACCTGTGAGCGGGACATC 914
QY 222 AlaIeTyThrIleSerProTyPheGlyThrIleGlySerValGluValTyr 240
Db 915 CTTGGCAGGTCTCATTTTGCGGCTCACAGTGTGCGCACCTGG 959
QY 241 LeuSpAspPheLeuLeuLeuMetLeuGlyIleProTyr 257
Db 960 TGT TGGTGACACAGACCGATCAT 983
QY 258 PheGlnArgValIleuSerSerSerAlaThrTyAlaGlnValLeuSerPheLeuAla 277
Db 984 GTGCGCGGTCTCTCGGCCAAGTCTGTCTCATGCGAAGGAGGCTCGTGTGGGG 1043
QY 278 AlaPhe 285
Db 1044 GGCTACCTGAAGATCTCCCATGTTCTTCATGTCGTCGCGGATATATCAGCGGGCC 1103
QY 286 Leu 296
Db 1104 CTGTCCACAGAGAGTGGCGCTGCGAGACCTGATGTCGTCGCAAGATCTGTGGGCGC 1163
QY 297 SerThrAspTyrAsnGlnThrAlaTyThrIleTyrProAspProIleThrIleGluIleAla 316
Db 1164 CGAGTGGGATGTTCCAACTTCCTAC 1190
QY 317 AspMetIleuProIleValIleuGlnTyIleuCySProValTyrIleSerPhePheGly 336
Db 1191 CTTAAGTGGTCAAGGCCCTCATGCTGTGTGCTGGGGGCGTATG 1238
QY 337 LeuGlyAlaValSerAlaAlaValMetSerSerAlaAspSerSerIleuSerAlaSer 356
Db 1239 ATTGCCGTGATCAGCGCTCTCATGAGCTCCTCATCTCATCTTCAACGAGCGAGC 1298
QY 357 SerMetPheAlaArgAsnIleTyThrIleuSerPheArgGlnAsnAlaSerAspIleGlu 376
Db 1299 ACCCTTTCACCATTCATGATGTGCGAG--CGCTTCCGAGAAAGTCAACAGAGAGAG 1355
QY 377 IleValTyrValMetArgIleThrValIleuValPheGlyAlaSerAlaThrAlaMetAla 396
Db 1356 CTGATGCGTGGCGAGAGTGTGTG--GTGTTC 1388
QY 397 LeuLeuThrIleValTyThrIleTyr 406
Db 1389 CTGGTTCATCAGGATC--CTGTGATCCCATTCACAAAGCTCAACAGTGGG 1442
QY 407 TyLeuSerSerAspLeuValTyIleIle 11lePheProGln 420
Db 1443 CAGCTTTCGACTCATTCAGGCTGTACCAAGTACCTGGCCCAACCATTCACCGCTTC 1502
QY 421 LeuLeuCyValLeuPheIleTyThrAsnThrTyThrIleValAlaIleGlyIleTyIle 440
Db 1503 TTCTGCTGGCGCATCTTCTGCMAAGAGGTCACAGAGCCGGAGCTTTCTGGGGCGCTCGTG 1562

QY 441 PheGlyLeu 443
Db 1563 TTTGGCGCTG 1571

RESULT 7
US-11-021-837-39/C
; Sequence 39, Application US/11021837
; Publication No. US20060140972A1
; GENERAL INFORMATION:
; APPLICANT: Manning, Paul
; APPLICANT: Alm, Richard
; APPLICANT: McLaughlin, Robert
; APPLICANT: McCormack, Kathleen
; TITLE OF INVENTION: Staphylococcus Saprophyticus Nucleic Acids and Polypeptides
; FILE REFERENCE: 101324-US
; CURRENT APPLICATION NUMBER: US/11/021, 837
; PRIOR FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/533534
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/600680
; PRIOR FILING DATE: 2003-08-31
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39
; LENGTH: 95963
; TYPE: DNA
; ORGANISM: Staphylococcus saprophyticus
US-11-021-837-39

Alignment Scores:
Pred. No.: 4,98e-13 Length: 95963
Score: 235.00 Matches: 124
Percent Similarity: 40.1% Conservative: 102
Best Local Similarity: 22.0% Mismatches: 249
Query Match: 7.9% Indels: 88
DB: 7 Gaps: 21

US-10-724-806-4 (1-580) x US-11-021-837-39 (1-95963)

QY 9 ValAlaIleIleLeuPheTyThrIleuIlePheLeuValGlyIleTyrAlaIleTyrIle 28
Db 55486 ATTAATGATCATCATATATATTTACATATATATTTATTTATGCTTACAGCA-----TAC 55433
QY 29 ThrIleAsnSerGlyAsnProGluIleArgSerGluAlaIleIleValGlyIleArgAsp 48
Db 55432 AAACAGCTACAGGTAATTAAGTGA-----TTTATGTTAGGTGAGAGC 55385
QY 49 IleGlyLeuLeuValGlyIlePheThrMetThrAlaThrTyrValGlyIleTyIle 68
Db 55384 ATCGTCTTATGTGACAGCACTTTCACAGAGTGCATGATAGTGGCTGATGATT 55325
QY 69 AsnGlyThrAlaGluAlaValTyThrIleProGlyCysGlyLeuAlaTyrAlaHisAlaPro 88
Db 55324 ATGGGTCTACACAGATCTGTTATAGTCTGGTATATGCTATGCTATGCTG-----ATTACC 55271
QY 89 IleGlyTyThrSerIleuIleLeuGlyIleuPhePheAlaIlePheMetArgSer 108
Db 55270 ATGTGCTTTCAATAGGCGCTATGTAAT-----TATTTGTGCTGTCACCTGTGTA 55217
QY 109 LysGlyTy 121
Db 55216 CGTGTCTATACGAATTTGGCTGGATGCGCATTCATTCACAGATTTCTTTAAAAACAGA 55157
QY 122 TyThrGlyIleArgMetGlyIleLeuPheIleProAlaLeuMetGlyIleMetPheTyr 141
Db 55156 TTAATATGACCAATTAATTAATTAATAAATTAATCCGCTTAATTAATGTTGTTTTC 55097
QY 142 AlaAlaIleIlePheSerAlaLeu-----GlyAlaThrIleSerValIleIleAsp 158
Db 55096 ACTTTATACACACATTCGGGTTTGTATCAGGTGTAATTAATTCGAAAGTCCCTCGGT 55037
QY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyThrIleVal 178
Db 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyThrIleVal 178

Db 55036 TTGAATTATCATTTGGGATTAATGCTTGTGATTTATAGTCATATTATTAACATTTTTC 54977
 Qy 179 GYGLYLeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGly 198
 Db 54976 GGTGGTATCTTCTGATCTATCATCACCAGCTTCTTCCAAAGAGTATCATGTTAAAGCGCT 54917
 Qy 199 LeuTyrPheSerValProPheAlaLeu-----SerHisPro 210
 Db 54916 ATG---GTATGGTCCCTTATTCGTTCGACCTTATGATTTTAAAGGCATTCATGATTTTAA 54860
 Qy 211 AlaValThrAspIleGlyPheThrAlaValHisAlaLeuTyrGlnSerProTyrPheGly 230
 Db 54859 CAAGTGCAGAGATGAACCAACCAAAATATGATCTTTTAAAGCACAACCGTTTATAGAT 54800
 Qy 231 ThrIleGluSerValGluValTyrThrTyrPheLeuAspAsnPheLeuLeuMetLeuGly 250
 Db 54799 ATCAAT-----TCACGTGTTGCTTG-----GGATTGGGT 54770
 Qy 251 GlyIleProTyrGlnAlaTyrPheGlnArgValLeuSerSerSerSerAlaThrTyrAla 270
 Db 54769 TATTTGCGAACCTCATATTATTTGTTAGTTTATGCTATAAATCGCAAAATTATTA 54710
 Qy 271 GlnValLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAlaLeuProAlaIleCys 290
 Db 54709 CCAAAAGCAGCAAGATTAGGTATTAGCTGATGATGATGCTTATAGGCTGTTGCA 54650
 Qy 291 IleGlyAlaIleGlyAlaSerThrAspTyrAsnGlnThrAlaTyrGlyTyrProAspPro 310
 Db 54649 GTTGGTTTAAACAGATTTGGCTTT---ATTTCTGAACGCAACATCAAGTTAGAGATCCCT 54593
 Qy 311 LysThrLysGluGlnIleAspMetIleLeuProIleValLeuGlnTyrLeuCysProVal 330
 Db 54592 GAAACT-----TTATTATGCTTATGACCAAAATTTATTCACCCCT 54551
 Qy 331 TyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaValMetSerSerAlaAspSer 350
 Db 54550 TTAGTTGGAGGCTTCTTATTTGGCAGCAATTTTGGCAGCCATCATGATGATCAATTTCTTCA 54491
 Qy 351 SerIleLeuSerAlaSerMetPheAlaArgAsnIleTyrGlnLeu-----Ser 367
 Db 54490 CAATTACTGTAACTTCAAGTCTCTTAACGGAAGACTTTTAAACCTTTCGTGCTGAA 54431
 Qy 368 PheArgGlnAsnAlaSerAspLysGluIleValTyrValMetArgIleThrValLeuVal 387
 Db 54430 GATTAAGCTTAAGCACAACGAGAAAGATCTTAATGTTGCTGCTTATCTGATTAATTT 54371
 Qy 388 PheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLysThrValTyrGlyLeuTyr 407
 Db 54370 GTAGCGATTGTGCTTATGATGATTTGCT-----TGGTCG 54338
 Qy 408 LeuSerSerAspLeuValTyrIleIle-----Ile 417
 Db 54337 CCAAAAGATACGATTAATTAATTTAGTTGATGCTTGGGAGAGATTGGTCCGCATTT 54278
 Qy 418 PheProGlnLeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAla 437
 Db 54277 AGTCCATTTGATATCTTCTCATCTCTATGGAAGGCTATACGTACTGCTGCACTTGA 54218
 Qy 438 GlyTyrIlePheGlyLeuPheLeuArgIleThrGlyGlyLeuProTyrLeuTyrLeuGln 457
 Db 54217 GGATGATTAACCGCCCATTTGCTGCTAATTTATTTGG-----ATTGTTGGATTTAA 54167
 Qy 456 ProLeuIlePheTyrProGlyTyrTyrSerAspLysAsnGlyTyrIleArgGlnArgPhe 477
 Db 54166 CCATTA-----GCTTCTATCAATGACCTATTTGCTATGAT---GAAATTAAT 54122
 Qy 478 ProPheLysThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeu 497
 Db 54121 CCTGGAATCTTACAGATGATGATTAATTAATTT-----GTT 54083
 Qy 498 AlaLysTyrLeuPheGluSerGlyThr---LeuProProLysLeuAspValPheAspAla 516
 Db 54082 AGTAATAATACATAAAACCTGGTCAATTTGTGACACATGACCTAGATAAAACAA 54023

Qy 517 ValValAlaArgHisSerGluGlnLeuMetAspLysThrIleLeuValArgAsnGlnAsn 536
 Db 54022 ATCGTA-----AAGGAATAGTTGATATATTAAGCTGTTAAATCAATACAAAC 53975
 Qy 537 IleLysLeu 539
 Db 53974 GCCAAACTG 53966
 RESULT 8
 US-11-217-529-1082
 / Sequence 1082, Application US/11217529
 / Publication No. US20060099612A1
 / GENERAL INFORMATION:
 / APPLICANT: SUNTORY LIMITED
 / APPLICANT: NAKAO, YOSHIHIRO
 / APPLICANT: NAKAMURA, NORIHISA
 / APPLICANT: KODAMA, YUKIKO
 / APPLICANT: FUJIMURA, TOMOKO
 / APPLICANT: ASHIKARI, TOSHIHIRO
 / TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
 / FILE REFERENCE: S-38-285
 / CURRENT APPLICATION NUMBER: US/11/217,529
 / PRIOR FILING DATE: 2005-09-02
 / PRIOR APPLICATION NUMBER: US 10/932,182
 / NUMBER OF SEQ ID NOS: 197023
 / SOFTWARE: PatentIn version 3.3
 / SEQ ID NO 1082
 / LENGTH: 2211
 / TYPE: DNA
 / ORGANISM: Saccharomyces pastorianus
 US-11-217-529-1082
 Alignment Scores:
 Pred. No.: 1,95e-12 Length: 2211
 Score: 206.50 Matches: 126
 Percent Similarity: 36.3% Conservative: 87
 Best Local Similarity: 21.5% Mismatches: 245
 Query Match: 6.9% Indels: 129
 Gaps: 23
 US-10-724-806-4 (1-580) x US-11-217-529-1082 (1-2211)
 Qy 39 SerGluAlaIleIleValGlyArgAspIle-----GlyLeuLeuValGlyGlyPhe 56
 Db 133 GCTGAAGAAATTTACCACTGCCGTAGATCGGTAAAGACCGGTTACTGGCGCGCTGTG 192
 Qy 57 ThrMetThrAlaThrTyrValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyr 76
 Db 193 GTTTCAGTTGATTTGGTGTGCTTACGTTGACTTCACTTCCAAAGAAATGCGAGAT 252
 Qy 77 GlyProGlyCysGlyLeuAlaTyrPalahisAlaProIleGlyTyrSerLeuSerLeuIle 96
 Db 253 GGATATTTGGTGGGACCAATATGCC-----GCCGTGCTTGTTCCTCAAAATTAAT 303
 Qy 97 LeuGlyGlyLeuPhePheAlaLysProMetArg---SerLysGlyTyrValThrMetLeu 115
 Db 304 GCGTTGCCATCTTGGCTTATTAAGACCAAGCAATGCTCCCAATCGCACACATCTTG 363
 Qy 116 AspProPheLysGlnIleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeu 135
 Db 364 GAATTAGTAGAAGAAAAGATATGATGATGGCCATGATGCTATGCTTTACGCCAT 423
 Qy 136 MetGlyGluMetPheThrPalahisAlaIlePheSerAlaLeuGlyAlaThrIleSerVal 155
 Db 424 GCAACCAATTTTGGTCACTTATGCTTTTAACCTTCAAGGCTCCGCTGTTTCAAGTAC 483
 Qy 156 IleIleAspValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyr 175
 Db 484 TTGACCGGATGAAACCACTTGATCTTCTTCTTCACTACCGCTGCTGTGCTTAT 543
 Qy 176 ThrLeuValGlyGlyLeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIle 195

```
Db 544 ACTCTGTTGGTGGATGATCAAGGCACTTTTCTAACCGACTATATTCACACTTGTGTCATC 603
Qy 196 PheIleGlyLeuTrpIleSerValProPheAlaLeuSerHisProAlaValThrAspIle 215
Db 604 ATCATC-----ATTGTTCTCGTATTTGCACTTTAAAGTTTAAAGTTCACCTAGTATATT 654
Qy 215 ----- 215
Db 655 TTAGGTGCGCGGCAAGTTTATGACTTATGCTGAGGCTGCTAAGACACCCAGTC 714
Qy 216 -----GlyPheThrAlaValHisAlaLysTyrGlnSerProTrpLeuGlyThrIleGlu 233
Db 715 GACGGTAATTTGAAGGTGAGTATGACTATGACATGACATAAATCCCGGTAATTTTATG 774
Qy 234 SerValGluValTyrThrTrpLeuAspPheLeuLeuLeuMetLeuGlyGlyIlePro 253
Db 775 ATCATTAATCTG-----ATCGGAAATTTCCGTTACCGTTTCTTGAT----- 816
Qy 254 TrpGlnAlaTyrPheGlnArgValLeuSerSerSerSerAlaThr-----TyrAlaGln 271
Db 817 ---AACGGTTATGGAATAAGGCCATTTCTGCTAGTCTCCGACGAAGTTTGAAGCTTAC 873
Qy 272 ValLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIle 291
Db 874 GCCATGGGTGGTGAACATGGTTGCT-----GTGGCTTCTCTTATTTGATGACACATG 927
Qy 292 GlyAlaIleGlyAlaSerThrAspTrpAsnGlnThrAlaTyrGlyTyrProAspProLys 311
Db 928 GGTTAGCATGATTCGACGTGAGACCTCTCCAAATTTCCCACTATCTGATCTCATCACTA 987
Qy 312 ThrIysGluGluAlaAsp-----MetIleLeuProIleValLeuGlnTyrLeuCysPro 329
Db 988 ACTTCGTTCACGGCAAAATTCGGGTGGTTGGTTTGGCCACGCTGCCGCA----- 1032
Qy 330 ValTyrIleSerPheGlyLeuGlyAlaValSerAla----- 342
Db 1033 -----ATTGCCATCATGGTGAAGAGGTGCAGTGGCATCCCTGCTGATCTTCATG 1086
Qy 343 AlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsn 362
Db 1087 GCCGTCACTGTGCCATGCTGCTGAATTGATGGCGTGTACTGTTTCACTTACATCAAT 1146
Qy 363 IleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValIleTyrValMetArg 382
Db 1147 ATCTATAGAGATATATATGATCTCTGTCAGACGGTAAGAACTGATTTACACCTCGAC 1206
Qy 383 IleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIysThrVal 402
Db 1207 TGTGCTGTATCTTTTGTGTTGCTTGCATGAGTGAATTTCTGTC----- 1251
Qy 403 TyrGlyLeuTyrTyr-----LeuSerSerAspLeuValTyr-----IleIle 416
Db 1252 ---GGACTGTACTATGTTGGTATTTCCATGGGCTATATTTATGAATGATGGTAAATTT 1308
Qy 417 IlePheProGlnLeuLeuCysVal-----LeuPheIleLysGlyThrAsnThrTyr 433
Db 1309 ATTAGTACTGACGTGTGCTGTGTTGACTTTGTGCTCTTAAAGACATGATGATGAT 1368
Qy 434 GlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArgIleThrGly----- 449
Db 1369 GCCGCTGTGTGTCCTCCCAATTTTAAAGCACAGCTTGGCTATATATGTCATGTTGGTTGC 1428
Qy 449 ----- 449
Db 1429 ACCAAGTCACTTACATGAACATACTGAGACACACATTCATGATATTCATTCATGCTA 1488
Qy 450 ---GlyGluProTyrLeuTyrLeuGlnProLeuIlePheTyrPro-----GlyTyrTyr 466
Db 1489 ACAGGTAAATTTGGTAGCTTTATTTGTCACACGCTATCTTATTTCCATTTTGAATTTTC 1548
Qy 467 SerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLysThrLeuSerMetValThr 486
```

```
Db 1549 -----TTTAAACCGCAAAATTTGCACTGGGAAAAATTCAGAGATTAC 1593
Qy 487 SerPheThrAsnIleCysValSerTyrLeuAlaLysTyrLeuPheGluSerGlyThr 506
Db 1594 AGA-----GTAGATGAATTCAGAA 1614
Qy 507 Leu-----ProProLysLeuAspValPheAspAlaVal--AlaArgHisSer 522
Db 1615 TTGGCTCAAGCTGATCCGATATTCAGATTTATGACGCTGAGCTTAATGCTTAAGGCTCA 1674
Qy 523 GluGluAsnMetCAspLysThrIleLeuValArgAsnGlu----- 535
Db 1675 GAGAGAAACAATTTCCCTGACTTTCGACAGAAATGAAAAAATGATGTTAGACTAAT 1734
Qy 536 AsnIleLysLeuAsnGluLeu 542
Db 1735 AATGAAAAATTTGAATGACCTC 1755
```

RESULT 9

```
US-11-217-529-76899
; Sequence 76899, Application US/11217529
; Publication No. US2006009612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217, 529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76899
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-76899
```

```
Alignment Scores:
Pred. No.: 2,51e-12 Length: 2208
Score: 205.50 Matches: 125
Percent Similarity: 38.3% Conservative: 101
Best Local Similarity: 21.2% Mismatches: 262
Query Match: 6.9% Indels: 103
DB: 8 Gaps: 21
```

```
US-10-724-806-4 (1-580) x US-11-217-529-76899 (1-2208)
Qy 39 SerGluAlaIleIleValGlyArgAspIle-----GlyLeuLeuValGlyLysPhe 56
Db 133 GCGAGCAATTTACCAACCGCGCGAGTCTGTGAAAAACCGGTATGTCGCCCGCTG 192
Qy 57 ThrMetThrAlaThrTrpValGlyGlyIleAsnGlyThrAlaGluAlaValTyr 76
Db 193 GTTCTAGTGGACATCTGTGTCTTACATGTTTAAAGTGTCAACAAAGAAATGACAGAC 252
Qy 77 GlyProGlyCysGlyLeuAlaTrpAlaHisAlaProIleGlyTyrSerLeuSerLeuIle 96
Db 253 GGTATATTGGCGGTATGCGGTACGCT-----GCTGGCGCATGCTTCCAAATTAAT 303
Qy 97 LeuGlyGlyLeuPhePheAlaLysProMetArg--SerLysGlyTyrValThrMetLeu 115
Db 304 GCATTCGCAATTTCCGCAATTAACCAAGCAAAATGGTCCCAATGCGCACACATATTTA 363
Qy 116 AspProPheLysGlnIleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeu 135
Db 364 GAATTAGAGAAACAAGATATGTAAGATGGCCATGGCTGCTACTGTTTATGCCATC 423
```


QY	89	IIegIyTyrSerLeuSerLeuIleuLeuGlyGlyLeu-----PhePhaIalysProMetArg	107
Db	253	GGTACATTCGATCGATCGTCGCAATTATTCATCTACTATTATTTATTCATGTCCTTTCTTT	312
QY	108	SerIySgIyTyrValThrMetLeuAspProPheIySgIle--TyrGly-----	123
Db	313	AAAAAGTTAAAGGTAAACATCTGCATATGATAATTTTAAAGAGTACATTTGGCCCTGCATCA	372
QY	124	LysAspMetGlyGlyLeuLeuPheIleProIalaleuMetGlyIleuMetPheTrpAlaIa	143
Db	373	CGTGCATTCGGCTCATTTATTTATTTGTCTGTTACCACTTTAGGCGCGTT-----GCAATT	426
QY	144	AlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAspValAsnIle	163
Db	427	GTTATCTACTTA-----CCAACATTAGCAATCAATCATCTGTATCAGACATGAACCT	477
QY	164	SerValIleValSer-----AlaIleuIleAlaIleuIyTrpThrLeuValGlyGlyLeu	181
Db	478	TATATCGTGCATCACTCGCTGGTTGGTTACTATGATATTTATTAATACATTTTATAGTGGTTTC	537
QY	182	TyrSerValAlaTyrThrAspValAlaGluLeuPheCysIlePheIleGlyLeuTrpIle	201
Db	538	GAAGGTGGTGTGGAGTGAATTTCATTCAGGCCATCTTATTATAGGGCGCTTTAGTT	597
QY	202	SerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaValHis	221
Db	598	-----ATTATATATCTAGGTGGTTGTGCACATTTAA	627
QY	222	AlaIySyrGlnSerProTrpLeuGlyThrIleGlnSerValGluValTyrThrTrpLeu	241
Db	628	GGCGGTTTCGGACACTGCTTTGCAATGCGATTGAGCACAAAAATTATTAAGT---GCA	684
QY	242	AspAsnPheLeuLeu-----LeuMetLeuGlyGlyIle	252
Db	685	GACAAATTGGAAACATTAATCTGCGGACGCTGCATTCAATTAATTTCTTGGAAAAATAT	744
QY	253	-----ProTrpGlnAlaTyrPheGlnArgValLeuSer	264
Db	745	TTCAACAACTGTATCAATPACACAGCGAGTCAAGACGTGTGCACGTTATCAACCTCT	804
QY	265	SerSerAlaThrTyrAlaGln-----ValLeuSerPheLeuAla	277
Db	805	GATACGTTTAAAGAAACAAATTAATCGTTATGACAAATGGATCTTACGTTTAAATTC	864
QY	278	Ala-----PheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIle	294
Db	865	GCACCTTATTTTAAAGTAAAGGTGCAATGCTGTATTCATTTTATPACACATGAACCTGT	924
QY	295	-----GlyAlaSerThrAspTrpAsnGlnThrAlaTyrGlyTyrProAspProIys	311
Db	925	TTACCAAAAGGCTTCAATCA-----	945
QY	312	ThrIySgIuGluAlaAspMetIleLeuProIleValLeuGlnTyrLeuCysProValTyr	331
Db	946	-----TCATCTGTAGTCGCAATTTTCATTTTGACATGAGATGCCACCATTT	990
QY	332	IleSerPhePheGlyLeuGlyAlaIalaserAlaIalaserAlaIalaserSerAlaAspSer	351
Db	991	GTAGGAGGATTACTTATTCAGCCATTTTGGCGCTGCACAGCTTACCATTTTCATCTAGT	1056
QY	352	IleLeuSerAlaSerSerMetPheAlaArgAsnIleIyGlnLeuSerPheArgIleAsn	371
Db	1051	TTAAATTTCTATCTGCTTGTGTTTTCATATGCACATTAAGCAACGCTCTTGGAAGAAAGT	1110
QY	372	AlaSerAspIySgIuIleValTrpValMetArgIleThrValLeuVal-----	387
Db	1111	AGCGAGCAGACAGAGTTAACTTGCACGTTTCATTATATTCATTCAGAGTAATTTTCGT	1170
QY	388	PheGlyAlaSerAlaThrAlaMetAlaLeuLeuTrpIySthValTyrGlyLeuTrpIyTr	407
Db	1171	TTTGGAAATGTCACTTACTTATTCGCTTCTTATTCAAATGACTTATGGGATTTATTTCTTG	1230
QY	408	LeuSerSerAspLeuValTyrIleIleIlePheProGlnLeuLeuCysValleuPheIle	427

Db	1231	TTTGTACCTGATTAATTCGGCGTTCACATTCGGCTGCTGATATTGCGATGGTATTTTCTACT	1290
Oy	428	LysGlyThrAsnThrYrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuAryIle	447
Db	1291	AAAGGTACGAATACATTCGGGTGATTAATTTGGGATTAATATTTGGGTATCATCTTT	1344
Oy	448	ThrGlyGlyIleuProTyrLeuIleuGlnProleuIlePheTyrProGlyTyrTyrSer	467
Db	1345	-----GGTTATGCTAT-----	1356
Oy	468	AspLysAsnGlyIleYrAsnGlnAryPheProPhe-----LysThrLeuSerMetVal	485
Db	1357	-----AATGGTGTGGCAAAAGGTAATCTCACTTTCTATGATTAATCACTTCACTTTTACA	1410
Oy	486	ThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyrLeuPheGluSerGly	505
Db	1411	GTTCCTTTTGCTTTGCTTCTTATATACCTTAAGCTTCATTT	1446
Oy	506	ThrIleuProPyrLysLeuAspValPheAspAlaValAlaArgHisSerGluGluAsn	525
Db	1447	-----GTCCCTTCAAAACATAAAAAAGATATA	1473
Oy	526	MetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsnGluLeuAlaProVal	545
Db	1474	ACGGGATTACCAATTTT-----GAA	1494
Oy	546	LysProArgIleSerLeuThrLeuSerSerThrPheThrAsnLys	560
Db	1495	AAAGATAAACCATCAACATACATTTCAAAAAACGGCTACGAAAAAG	1539
RESULT 11			
US-11-266-748A-26072			
Sequence 26072, Application US/11266748A			
Publication No. US20060134663A1			
GENERAL INFORMATION:			
APPLICANT: Hartkin, Paul			
APPLICANT: Johnston, Patrick			
APPLICANT: Mulligan, Karl			
TITLE OF INVENTION: Transcription Microarray Technology and			
TITLE OF INVENTION: Methods of Using the Same			
FILE REFERENCE: 55815-0102 (319189)			
CURRENT APPLICATION NUMBER: US/11/266, 748A			
CURRENT FILING DATE: 2005-11-03			
PRIOR APPLICATION NUMBER: EP 04105479. 2			
PRIOR FILING DATE: 2004-11-03			
PRIOR APPLICATION NUMBER: EP 04105482. 6			
PRIOR FILING DATE: 2004-11-03			
PRIOR APPLICATION NUMBER: EP 04105483. 4			
PRIOR FILING DATE: 2004-11-03			
PRIOR APPLICATION NUMBER: EP 04105507. 0			
PRIOR FILING DATE: 2004-11-03			
PRIOR APPLICATION NUMBER: EP 04105485. 9			
PRIOR FILING DATE: 2004-11-03			
PRIOR APPLICATION NUMBER: EP 04105484. 2			
PRIOR FILING DATE: 2004-11-03			
PRIOR APPLICATION NUMBER: US 60/662, 276			
PRIOR FILING DATE: 2005-03-14			
PRIOR APPLICATION NUMBER: US 60/700, 293			
PRIOR FILING DATE: 2005-07-18			
NUMBER OF SEQ ID NOS: 483996			
SOFTWARE: PatentIn version 3.3			
SEQ ID NO 26072			
LENGTH: 2984			
TYPE: DNA			
ORGANISM: Homo Sapiens			
US-11-266-748A-26072			
Alignment Scores:			
Pred. No.:	1,566-10	length:	2984
Score:	191.00	Matches:	120
Percent Similarity:	38.1%	Conservative:	70
Best Local Similarity:	24.0%	Mismatches:	197

Query Match:	6.4%	Indels:	112
DB:	8	Gaps:	24
US-10-724-806-4 (1-580) x US-11-266-748A-26072 (1-2984)			
Qy	3 PheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeuValGly	22	
Db	278 TTCTCCATCAGGACTATGTGTGTGCTGGCTGCGTGTCTCTCTCTGCAATTGGG	337	
Qy	23 IleTPrpAlaIaIa-----TPrpLysThrLysAsnSerGlyAsnProGluGluArgSer	39	
Db	338 CTCTACCATGCTGTCTGCGCTGGGCGCGCATCTGTGGTAG-----	382	
Qy	40 GluAlaIleIleValGlyLysArgAspIleGlyLeuValGlyGlyPheThMetThr	59	
Db	383 -----CTGCTGATGGCGCGACCGAATAATGGGCTGCTTCCGGTGGCATGTCTCTGCTG	436	
Qy	60 AlaThrTPrpValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGly-----	77	
Db	437 GCCACCTTCCAGTCAGCGCGTGGCCATCTGGGTGTGGCCGTACAGAGATCTACGATTGGG	496	
Qy	78 -----ProGlyCys-----GlyLeu-----AlaTPrpAla	85	
Db	497 ACCCAATATGTGTTCTCTGGCTGCTGCTACTTTCGGGGCTGTGATCTGCACACATC	556	
Qy	86 HisAlaProlIleGlyTyrSerSerLeuSerLeuIleGly-----GlyLeuPhe	101	
Db	557 TTCAATCCCGTTTCTTACCGCTGCTGATCTACACAGTGGCTATGAGTACCTGAGCTTGA	616	
Qy	102 PheAlaLysPProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGluIle	121	
Db	617 TTCAATATAAAGCTGCGCGAGTGTGTGAGACTGTGACCTTCAATC-----TTTCAAGATGGT	670	
Qy	122 TyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyLysMetThr	141	
Db	671 ATC-----TACATGGAGGATGTGTGCTCTATGCTCCGTGATTTG-----	706	
Qy	142 AlaAlaAlaIlePheSerAlaLeuGlyLahThrIleSerValIleIleAspValAspVal	161	
Db	707 -----GCTCTCAATGCAAGTACTGGCTTTGATGTG	736	
Qy	162 AsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGlyLeu	181	
Db	737 TGGGTGTGCGTGTGCGCCCTGGCATTTCTGTACCGCTTATACAGCTGTGGTGGCGTG	796	
Qy	182 TyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTPrpIle	201	
Db	797 AAGGCGCATCTGAGACAGATGTGTTCCAGACACGTGATGTTCTCTGGG-----	847	
Qy	202 SerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal---	220	
Db	848 -----CACGTGGCAATATCATGTGGGGGTACGCCAAGGTGGGC	886	
Qy	221 -----HisAlaLysTyrGlnSerProTPrpLeu	229	
Db	887 GCGTTGGGGCGCTGTGTGGCCGTGCTTCCAGACAGCGCCGACATCTGTGGTTAGAGCTG	946	
Qy	230 GlyThrIleGluSerValGluValTyrThrTPrpLeu-----AspAsnPhLeu	245	
Db	947 GATCCAGACCCCTTTGTCGGCACACCTTCTGACCTTGGCCCTTGGGGGTGTCTTTCATG	1006	
Qy	246 LeuLeuMetLeuGlyGlyIleProTPrpGlnAlaTyrPheGlnArgValLeuSerSerSer	265	
Db	1007 ATGCTCTCTTATACGGGGGTGAC---CAGGCTCAGGGTACAGGGTACTCAATTCCTCGGC	1063	
Qy	266 SerAlaThrTyrAlaGlnValLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAla	285	
Db	1064 ACGGAGAG--GCGTGTGTGCTCTCTGTTATGACAGTTCCTCCCTTCCAGACAGGTGCC	1120	
Qy	286 LeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTPrpAsnGlnThrAlaTyr	305	
Db	1121 CTC-----TGGTGGGCTGCTCATGTCCTGATC-----ATGTTCGGCGAT	1162	

Qy	306	GlyTyrProaspProLysThrLysGlnIleuAlaSerMetIleLeu---ProIleValIleu	324
		::: :::::	
Db	1163	TACCAGGAGATATCCCATGAGCATTCCAGAGGCTCAGGAGGCCACGACGATGTCCTCG	1222
Qy	325	GlnTyrLeuCySProValTyrLisSerPhePheGlyLeu---GlyAlaValSerAlaIa	343
		::: :::::	
Db	1223	TACTTTGTGATGATCTTCCTGAAAGGGCTCGCCAGGCTCGCCAGGGCTCTTCAATTCCTGCG	1282
Qy	344	ValMetSerSerAlaaspSerSerIleLeuSerAlaSerSerMetPheAla-----	360
		::: :::::	
Db	1283	CTCTTCAGGCGCTCTCCAGCATTATCTCTGCTTTTAATTCACTGGCAACTGTACG	1342
Qy	361	---ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrP	379
		::: :::::	
Db	1343	ATGGAAGACCTGATGATGACCTGTGCTCCCTGAGTCTCTGAAGCCGGGACATC-----	1396
Qy	380	ValMetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThr	399
		:::	
Db	1397	---ATGCTTTCACAGAGCCTTGCCCTTGGCTAGAGGGCTGTTGCTA-----	1441
Qy	400	LysThrValTyrGlyLeuTyrTyrLeuSerSerAspLeu-----ValTyr	414
		::: :::::	
Db	1442	-----GGATGGCTATATTTCTCCACAGATGGGACCTGTCTCGAGGACAGA	1489
Qy	415	IleIleIlePheProGln-----LeuLeuCysVal--LeuPhe	426
		:::	
Db	1490	ATCAGCAATCTTTGGCATGATGGGGAGACCGCTGCTGGACATCTTGCTTGGAATGTTC	1549
Qy	427	IleLysGlnLysAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlnLeuPheLeu	445
		:::	
Db	1550	TTTTCATGTCTAAACCTCTCGTGAGTGTGTGGGCTGTGGGCTGTGGCTGCTCATG	1606

```

RESULT 12
US-11-021-837-27/c
Sequence 27, Application US/11021837
Publication No. US20060140972A1
GENERAL INFORMATION:
APPLICANT: Alm, Richard
APPLICANT: Manning, Paul
APPLICANT: McLaughlin, Robert
APPLICANT: McCormack, Kathleen
TITLE OF INVENTION: Staphylococcus Saprophyticus Nucleic Acids and Polypeptides
FILE REFERENCE: 101324-US
CURRENT APPLICATION NUMBER: US/11/021,837
CURRENT FILING DATE: 2004-12-23
PRIORITY APPLICATION NUMBER: US 60/533534
PRIORITY FILING DATE: 2003-12-31
PRIORITY APPLICATION NUMBER: US 60/600680
PRIORITY FILING DATE: 2003-08-31
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn version 3.3
SEQ ID NO 27
LENGTH: 38840
TYPE: DNA
ORGANISM: Staphylococcus saprophyticus
US-11-021-837-27

Alignment Scores:
Pred. No.:          4.86e-08          Length:          38840
Score:              184.00           Matches:          118
Percent Similarity: 39.6%           Conservative:     100
Best Local Similarity: 21.5%        Mismatches:      222
Query Match:        6.1%            Indels:           110
DB:                 7                Gaps:             25
US-10-724-806-4 (1-5680) x US-11-021-837-27 (1-38840)
Oy      9 ValAlaIleIleuPheTyrlEuIeuIlePheIeuValGlyIleTrpAlaAlaTrpLys 28
      |||||:::|||||::: |||||::: |||||::: |||||::: |||||:::
Db      6204 GTACTGTAAATAGTTATCTCGTATTAATGATTATGGGGG-----GCCGATATTT 6154
Oy      29 ThrTysAsnErgGlyAsnProGluGlnArgSerGluAlaIleIleValGlyGlyArgAsp 48
      |||||::: |||||::: |||||::: |||||::: |||||::: |||||:::

```


Db	6155	ACAAAAACGAGCA-----AGTCAAAAACACAGATAGTTCTTTACACGACTAGTGGCCGT	6103
Qy	49	IIeGIyLeuLeuValGIyGlyPheThrMetThrAlaThrTPValGIyGlyTyTrIle	68
Db	6102	TTGGCTTCGTGGGCGAGTTGGGTTCCTTACCTATGACACACATAGCGCATTCATTT	6043
Qy	69	AeNGIyThAlaGIuAlaValTyrgIyProGIyCySGIyLeuAlaTPAlaHisIaPro	88
Db	6042	ATGTGACACCCAGAAAAGCATTTTAAACAGT-----TGGCTTACCTACT	5995S
Qy	89	IIeGIyTyrSerLeuSerLeuIIeIeuGIyGlyLeu--PhePhAlaIuysProMetArg	107
Db	5994	GGCAATATGCAATATGACTATATTCATATGTTAAATTAATTTATTAATGACATTCCTT	5935S
Qy	108	SeTySGIyTyValThrMetLeuAspProPheIySGinIle--TyrgIy-----	123
Db	5934	AAAAAATTAAAGGTGACATCAGCTTATGATATTTTAAAGACACGATTTGGGCCAGTCTC	5875S
Qy	124	LySArgMetGIyGlyLeuLeuPheIIeProAlaIeuMetGIyGluMetPheThrAlaIa	143
Db	5874	CCGTGATATGGGTGCTGTCTTATTTGATTTGATTCATATTAAGAAAGGTA-----GCTATT	5821S
Qy	144	AlaIIePheSerAlaIeuGIyAlaThrIIeSerValIIeIIaSpValAspValAsnIle	163
Db	5820	GTGATTTATTTA-----CCGACGCTAGGAGATTACTGCGGTTCTGATTTAAATCTT	5770S
Qy	164	SeValIIeValSer-----AlaIeuIIaIIaIIeLeuTyTrThLeuValGIyGlyLeu	181
Db	5769	TATGTGTCGACGAGCTACTAGAGAAATGTCTGATTTCTTATACATTCCTTGATGTTT	5710S
Qy	182	TyrSerValAlaTyTrThAspValValGInIeuPheCysIIePheIIeGIyLeuTrIle	201
Db	5709	GAAGTGTCGTATGGAGTCACTTCCTTCAAGGGCATCTCAATTAAGTGTGCT-----	5656S
Qy	202	SeValIProPheAlaIeuSerHisIProAlaValIThrAspIIeGIyPheThAlaValHis	221
Db	5655	-----GCACGAATCATTAICTTGGGATTT-----GTACAT	5626S
Qy	222	AlaIyTyrgInIserProTrieuGIyThrIle-----Glu	233
Db	5625	ATCAAA-----GATGTATGGAGAACTGATGACGACGCCCTGATCATCAATAAAA	5575S
Qy	234	SeValIeGIuValTyTrThTrieuAspAsn-----PheLeuLeuIeuMet	248
Db	5574	TTAAATCAGTCCAGATPATTTGAAAGTTTAATCTCCGCGGCGACCAATTCCTATTATTTTC	5515S
Qy	249	LeuGIyGlyIle-----ProTrieuAlaIyTyPheGInArg	260
Db	5514	TTAGGAAATATATTTAATTTTGCATCAATATACGCAAGTCAAGATGTTGTCACACGT	5455S
Qy	261	ValIeuSerSerSerSerAlaThTyrgIyArgInValIeuSerPheLeuAlaIaPheGly	280
Db	5454	TATCAAGCAATCGAATCCATGCTGTAAACGAAAAATCATTAATGATGAATAGTGTTC	5395S
Qy	281	CysLeuValMetAlaIeuProAlaIIeCysIIeGIyAlaIIeGIyAlaSerThAspTrp	300
Db	5394	GCATTTGATTTACGCCCATTTATTTATGTATAGTGAGACATCCTTATTCCTTACACA	5335S
Qy	301	AsnGIuThAlaTyTr-----GIyTyTrProAspProLySThrLySGIuAlaAspMet	318
Db	5334	CATGAAAGCTTTTCCGACAAAGTTT-----AATACATCATCG	5296S
Qy	319	IIeIeuProIIeValIeuGIuTyTrIeuCysProValTyTrIleSerPhePheGIyIeuGIy	338
Db	5295	ATTGTACTTATTTCATCTGACAGAAATGCCACCATTTATTTGCCGATTACTTAATTC	5236S
Qy	339	AlaValSerAlaIaValMetSerSerAlaAspSerSerIIeLeuSerValaSerSerMet	358
Db	5235	ACAAATATTTCGAGTCGACAAATCTCAATTCATCAAGCTTAATTAATTCATTTACGCTT	5176S
Qy	359	PheAlaIaArgAsnIIeTyrgInIeuSerPheArgGInAsnAlaSerAspLySGIuIleVal	378
Db	5175	TTATCTTCGAATATTAACATAGATTTTTCGCTAAAGGTAAGAAAGACAAAGATTTT	5116S

Qy	379	-----TPValMeArgIleThrValIleVal---	PheGlyAlaSerAlaIleThraIa	394
Db	5115	TTTGCTAGATGGATGATATATACCTTCGCTTATTTGGATTTGGTATGCTATCTACTTA		5056
Qy	395	MetAlaIeulThrIlyThsThValTyrgIlyIeuTrpTyIeuserSerAspLeuValTy		414
Db	5055	AMTGGTTCAGATCTCAATAGACTTATAGGATTTATCTCATATTAATGCGGCTGTAGGT		4996
Qy	415	IleIleIlePhePrGInIleuIeIeCysValIeuhelIeIysGlyThrAsnThrTyGly		434
		::: ::::	::: ::::	
Db	4995	GGCCATTGGCGGGGTATTTGCTGTAGATCTTACGAAACACAACAAATACCTTAGGC		4936
Qy	435	AlaValAlaGlyTyIlePheGlyIleuPheIeuArgIleThrGlyGlyIuPTrTyLeu		454
		::: ::::	::: ::::	
Db	4935	GTTATATCGGCTTTGGCGCTTAGGATATATCTTC-----GCTTATATC		4894
Qy	455	TyIeIeuGInProlIeulPheTyIreProGlyTyIreTyIserAspIysAsnGlyIleTyAsn		474
		::: ::::	::: ::::	
Db	4893	TTTC-----LeuAspValPheAsp		4876
Qy	475	GlnArgPhePrOphelyThIeuserMetValThrSerPhePheThraAsnIleCysVal		494
Db	4875	GGCAATTCGCCA-----TTCCTTGTATCCATT--ATT		4846
Qy	495	SerTyIeulAlaIysTyIrePheGlu-----SerGlyThrIeuProIlys---		510
Db	4845	TCATTCCATAGTGGCATTTATCTTCTCTTATCTATCAGTATGTATATCCTCGTAAGAA		4786
Qy	511	-----LeuAspValPheAsp		515
Db	4785	AAAGATATTGTTGGCTTAACATATTTTGAT		4756
RESULT 13				
	US-11-266-748A-27661			
	/ Sequence 27661, Application US/11266748A			
	/ Publication No. US2006013463A1			
	/ GENERAL INFORMATION:			
	/ APPLICANT: Harkin, Paul			
	/ APPLICANT: Johnston, Patrick			
	/ TITLE OF INVENTION: Transcriptome Microarray Technology and			
	/ FILE REFERENCE: 55815-0102 (319189)			
	/ CURRENT APPLICATION NUMBER: US/11/266,748A			
	/ CURRENT FILING DATE: 2005-11-03			
	/ PRIOR APPLICATION NUMBER: EP 04105479.2			
	/ PRIOR FILING DATE: 2004-11-03			
	/ PRIOR APPLICATION NUMBER: EP 04105482.6			
	/ PRIOR FILING DATE: 2004-11-03			
	/ PRIOR APPLICATION NUMBER: EP 04105483.4			
	/ PRIOR FILING DATE: 2004-11-03			
	/ PRIOR APPLICATION NUMBER: EP 04105507.0			
	/ PRIOR FILING DATE: 2004-11-03			
	/ PRIOR APPLICATION NUMBER: EP 04105485.9			
	/ PRIOR FILING DATE: 2004-11-03			
	/ PRIOR APPLICATION NUMBER: EP 04105484.2			
	/ PRIOR FILING DATE: 2004-11-03			
	/ PRIOR APPLICATION NUMBER: US 60/662,276			
	/ PRIOR FILING DATE: 2005-03-14			
	/ PRIOR APPLICATION NUMBER: US 60/700,293			
	/ PRIOR FILING DATE: 2005-07-18			
	/ NUMBER OF SEQ ID NOS: 483996			
	/ SOFTWARE: PatentIn version 3.3			
	/ SEQ ID NO 27661			
	/ LENGTH: 3844			
	/ TYPE: DNA			
	/ ORGANISM: Homo Sapiens			
	US-11-266-748A-27661			
Alignment Scores:				
	6e-08	Length:	3844	
	169.00	Matches:	139	
	Score:			


```
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 16097
/ LENGTH: 936
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-16097
```

```
Alignment Scores:
Pred. No.: 4,24e-06 Length: 936
Score: 143.50 Matches: 58
Percent Similarity: 36.8% Conservative: 37
Best Local Similarity: 22.5% Mismatches: 78
Query Match: 4.8% Indels: 85
DB: 8 Gaps: 9
```

US-10-724-806-4 (1-580) x US-11-266-748A-16097 (1-936)

```
QY 5 ValGluGlyLeuValAlaIleIleLeuPheTYrLeuLeuIlePheLeuValGlyIleTyr 24
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 70 GTGGCTGACATCATCGTCATCATCTGTGATTTTGTGATGTGGCCGTCATATG 129
QY 25 AlaAlaIleTyrLys---ThryLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIle 43
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 130 TCCTCTGTGGGCCGAGTGAACACGGTGAAT-----GGCTACTTC 171
QY 44 ValGluGlyArgAspIle----- 49
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 172 CTGGCAGGCGCGGACATGACGTGGCCGATTTGAGCCTCTTCGCGCAGCAGCAG 231
QY 50 -----GlyLeuLeuValGly 54
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 232 GGCTGTGGCTCTTCATTTGAGTGGCGGCTGACGCGGCGGAGAGTCTGGCCGTGGCA 291
QY 55 GlyPheThMetThrAlaThrTyrValGlyGlyGlyTyrIleAsnGlyThrAlaGluAla 74
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 292 GGCTTCGAGTGAATGCCAGTACGTG----- 318
QY 75 ValTYrGlyProGlyCysGlyLeuAlaTyrAlaIleAlaProIleGlyTyrSerLeuSer 94
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 319 -----CTGCTGGCAGCTGGCATGGGTGTGTGCTCCATCTACATCTCTCAGAG 366
QY 95 LeuIleLeuGlyGlyLeuPheAlaLys-----ProMetArgSerLysGlyTyr 111
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 367 ATCGTCACCTTACCTGAGTACATTTCAGAACGCTACGCGGCGGCGGAGTCCGATCTAC 426
QY 112 ValThMetLeuAspProPheLysGlnIleTyrGlyLys-----ArgMetGly 127
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 427 CTGTCTGTCTGTCTCTGCTGCTGCTGTGTACCAAGATATCGCTGACCTGTACGCG 486
QY 128 GlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPheTyrAlaAlaIlePheSer 147
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 487 GGGGCTCTGTGTGTGACATCTGCTGGGCTGGAACTTCTAC----- 528
QY 148 AlaLeuGlyAlaThrIleSerValIleIleAspValAspValAsnIleSerValIleVal 167
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 529 -----CTCTCCACCATCTCTC 543
```

```
QY 168 SerAlaLeuIleAlaIleLeuTYrThrLeuValGlyGlyLeuTyrSerValAlaIleTyrThr 187
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 544 ACGCTCGCATACACAGCCTGTATACACATCGCAGGTGGCTGTGCTGTATCTACAG 603
QY 188 AspValAlaGlnLeuPheCysIlePheIleGlyLeuTyrIleSerValProPheAlaLeu 207
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 604 GAGCCCTGCAGACGCTCATCATGGTGGTGGGGCTGTCATC-----CTG 648
QY 208 SerHisProAlaValThrAspIle---GlyPheThrAlaValHisAlaLysTyr 224
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 649 ACAATCAAGCTTTTGACCAATCGGTGGTACGGGCAAGCTGAGGCAAGCTTAC 702
```

Search completed: July 10, 2006, 14:56:38
Job time : 503 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:34 ; Search time 8716 Seconds
(without alignments)
11182.583 Million cell updates/sec

Title: US-10-724-806-3
Perfect score: 1743
Sequence: 1 atgccttcacatgctgaag.....ctgaataattacaatga 1743

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc3:*
3: gb_esc4:*
4: gb_esc5:*
5: gb_esc6:*
6: gb_esc7:*
7: gb_esc8:*
8: gb_esc9:*
9: gb_esc10:*
10: gb_esc11:*
11: gb_esc12:*
12: gb_esc13:*
13: gb_esc14:*
14: gb_esc15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1731.8	99.4	4097	6	AK053063	AK053063 Mus muscu
2	1731.8	99.4	4306	6	AK034415	AK034415 Mus muscu
3	1465	84.1	1743	14	AY413300	AY413300 Homo sapi
4	1162.6	66.7	1743	14	AY413298	AY413298 Homo sapi
5	1162.2	66.7	1743	14	AY413299	AY413299 Pan trogl
6	851.8	48.9	4122	6	AK164116	AK164116 Mus muscu
7	603.4	34.6	707	5	CD350164	CD350164 UT-M-FY0-
8	576.4	33.1	669	4	BY727598	BY727598 BY727598
9	576.4	33.1	882	5	CK704440	CK704440 958613 MA
10	472	27.1	828	14	CT403380	CT403380 Sus scrofa
11	409.6	23.5	650	7	BB626260	BB626260 BB626260
12	406.8	22.3	672	14	AG157499	AG157499 Pan trogl
13	398.4	22.9	548	9	DA361315	DA361315 DA361315
14	391.8	22.5	576	7	DN991044	DN991044 TC112868
15	389.6	22.4	516	7	BE233479	BE233479 139685 MA
16	370	21.2	675	4	BY729567	BY729567 BY729567
17	367	20.1	988	10	DM665738	DM665738 CNB333-A0
18	359.6	20.6	781	10	DT197042	DT197042 JGI_CAA55
19	359.6	20.6	843	10	DT207246	DT207246 JGI_CAA51

20	356.4	20.4	1114	10	DT972151	DT972151 CLJ138-H0
21	348.8	20.0	753	10	DV838598	DV838598 LB01121.C
22	325.6	18.7	874	5	CK449221	CK449221 892772 MA
23	302.2	17.3	834	9	CX840109	CX840109 JGI_CAA9
24	301.6	17.3	541	7	AM668962	AM668962 111664 MA
25	301.6	17.3	1054	7	DM614674	DM614674 CLJ282-B0
26	273	15.7	694	8	CX033587	CX033587 1345948 N
27	263.8	15.1	524	7	BE723927	BE723927 198406 MA
28	252.6	14.5	800	1	AL669749	AL669749 AL669749
29	251.2	14.4	652	7	BB626456	BB626456 BB626456
30	236.2	13.6	658	8	CX196909	CX196909 SA mxo 52
31	235.8	13.5	605	3	BQ829470	BQ829470 LUG12176
32	229	13.1	549	9	BM274870	BM274870 BM274870
33	228.8	13.1	1021	10	DM616805	DM616805 CLJ294-F1
34	226	13.0	504	5	CJ145625	CJ145625 CJ145625
35	225.4	12.9	941	5	CD360297	CD360297 AGENCOURT
36	220.6	12.7	565	9	DA911919	DA911919 DA911919
37	215.8	12.4	658	3	BM629925	BM629925 170006875
38	215.4	12.4	641	2	BI630566	BI630566 RH59836.5
39	215	12.3	574	9	DA366634	DA366634 DA366634
40	213.6	12.3	640	2	BI629504	BI629504 RH58381.5
41	212.8	12.2	598	8	CX196517	CX196517 SA mxo 46
42	212.4	12.2	1037	1	AL666817	AL666817 AL666817
43	209.8	12.0	682	9	DN604180	DN604180 ACAC-aab8
44	203.6	11.7	869	9	CX809396	CX809396 JGI_CAAJ1
45	200	11.5	583	3	BM277281	BM277281 BM277281

ALIGNMENTS

RESULT 1	AK053063	4097 bp	mRNA	linear	HTC 02-SEP-2005
LOCUS	AK053063	Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched library, clone:DJ0038E20 product:solute carrier family 5 (cholesterol transporter), member 7, full insert sequence.			
DEFINITION	AK053063	HTC, CAP trapper.			
ACCESSION	AK053063	AK053063.1 GI:26343192			
VERSION	AK053063	AK053063			
KEYWORDS	HTC, CAP trapper.	Mus musculus (house mouse)			
SOURCE	Mus musculus				
ORGANISM	Mus musculus				
REFERENCE					
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	10349636				
REFERENCE					
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	11042159				
REFERENCE					
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.				
TITLE	RIKEN Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
PUBMED	11076861				
REFERENCE					
AUTHORS	THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				

JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	
AUTHORS	5 The FANTOM Consortium, the RIKEN Genome Exploration Research Group phase I and II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6
AUTHORS	RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.
TITLE	Antisense transcription in the Mammalian Transcriptome
JOURNAL	Science 309, 1564-1566 (2005)
REFERENCE	7
AUTHORS	The FANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).
TITLE	The Transcriptional Landscape of the Mammalian Genome
JOURNAL	Science 309, 1559-1563 (2005)
REFERENCE	8 (bases 1 to 4097)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carrinci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hizomoto,K., Hirooka,T., Hirozane,T., Horii,F., Imocani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Karch,H., Kawai,J., Kohji,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishii,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sugabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takeda-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/. Location/Qualifiers
FEATURES	source 1..4097 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:D930038B20" /db_xref="taxon:10090" /clone="D930038B20" /tissue.type="head" /clone.lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="15 days embryo" 512..2254 /note="unnamed protein product; putative solute carrier family 5 (choline transporter), member 7 [MGI McI:1927126 GB NM_022025, evidence: BLASTN, 99%, match=1743]" /codon.start=1 /protein_id="BAC35253.1" /db_xref="gi:26345193" /translation="MSFRVEGLVAILLFYLLIPLVIGIAWAKTNSGNPERSBAIVAGGRDIGLVGGFTATVWGCGINGTAEAYGCGGLAWQAEIGSLILGLEFFAKPMRSGVTMLDPFOOITYGRMGGLFIYALNGEMMAAITSALGATSIIIDVDVNISYSALAIILTYTLGVGLSVAYDVOLFCIFIMKGYSPALSPAVDTIGFTAVAHKYOSPWICGTESVRYVWLDLNLMLGVLPGPKMXAPRVYSSSATSAQYLSTLAARCLWALPAITCIGAISTDMNQTRAGYDPDKTESEAMILLPIVQYICPVYSISFGDAVASVNVSSADSSITLSASWFARNYVTSFRQNASKEIVWMKITLVDFGASATMALMTKVYGLMWYLSDDIVYIIIFPOLCVLPFKIGNTVGA VAGYIFGLFIRTAG
CDS	

[illegible]

Oy 961 CCGATGCTTCGACGAGTACCTTCGCTGCTGATCTCTTCTTGGGGCTGGTCTGT 1020
 Db 1472 CCGATGCTTCGACGAGTACCTTCGCTGCTGATCTCTTCTTGGGGCTGGTCTGT 1531
 Oy 1021 TAGGCTGCTGATGTCCTGACGCTGATGTCCTGCTGGCGAGTTCTATGTTGCT 1080
 Db 1533 TAGGCTGCTGATGTCCTGACGCTGATGTCCTGCTGGCGAGTTCTATGTTGCT 1591
 Oy 1081 CGGAATATCTACGAGTTCTTCGACGAAATGATCGACGAAAGAAATGTTGGGTC 1140
 Db 1592 CGGAATATCTACGAGTTCTTCGACGAAATGATCGACGAAAGAAATGTTGGGTC 1651
 Oy 1141 ATGAGGATCATGTCGCTGCTGCTGAGATCTGCAACAGCCATGCTTGTGACGAAG 1200
 Db 1652 ATGAGGATCATGTCGCTGCTGCTGAGATCTGCAACAGCCATGCTTGTGACGAAG 1711
 Oy 1201 ACTGCTATGGGCTGCTGATCTGACCTGCTGATCTGCTCAATCATCTTCCACAG 1260
 Db 1712 ACTGCTATGGGCTGCTGATCTGACCTGCTGATCTGCTCAATCATCTTCCACAG 1771
 Oy 1261 CTGCTCTGTGTACTCTTCATCAAGGACCAACACTTATGGGCGAGTCTGTTATAT 1320
 Db 1772 CTGCTCTGTGTACTCTTCATCAAGGACCAACACTTATGGGCGAGTCTGTTATAT 1831
 Oy 1321 TTTGCACTATCTCTGAGAAATTAAGAGAGAGCCATATCTATCTTGGAGCCCTTAATC 1380
 Db 1832 TTTGCACTATCTCTGAGAAATTAAGAGAGAGCCATATCTATCTTGGAGCCCTTAATC 1891
 Oy 1381 TTCTACCTGCTGTTATCTCTGACAAAGATGTAATGTAATGTAATGTAATGTAAT 1440
 Db 1892 TTCTACCTGCTGTTATCTCTGACAAAGATGTAATGTAATGTAATGTAATGTAAT 1951
 Oy 1441 ACTCTCTCATGCTTACCTCTTCTTACCAAGATGTTGTTCTTATGACGAGTAT 1500
 Db 1952 ACTCTCTCATGCTTACCTCTTCTTACCAAGATGTTGTTCTTATGACGAGTAT 2011
 Oy 1501 CTATTGGAAGTGAACCTTGCTCCCAAAATTGATGTAATTTGATGCTGTCGCAAG 1560
 Db 2012 CTATTGGAAGTGAACCTTGCTCCCAAAATTGATGTAATTTGATGCTGTCGCAAG 2071
 Oy 1561 CACAGTGAAGAGAACCTGACCAAGACATTTCTAGTGAATGAATGAATGAATGAAT 1620
 Db 2072 CACAGTGAAGAGAACCTGACCAAGACATTTCTAGTGAATGAATGAATGAATGAAT 2131
 Oy 1621 GAACCTGACCTGTAACCTTGCGGAGACCTTACCTGATCACTTCACTCAATTAAG 1680
 Db 2132 GAACCTGACCTGTAACCTTGCGGAGACCTTACCTGATCACTTCACTCAATTAAG 2191
 Oy 1681 GAGGCGCTCTGATGTTGATTCAGTCCGAGAGGGCTGCGAGTGAAGTAATTTACAA 1740
 Db 2192 GAGGCGCTCTGATGTTGATTCAGTCCGAGAGGGCTGCGAGTGAAGTAATTTACAA 2251
 Oy 1741 TGA 1743
 Db 2252 TGA 2254

RESULT 2
 AK034415 4306 bp mRNA linear HTC 02-SEP-2005
 LOCUS
 DEFINITION Mus musculus adult male diencephalon CDNA, RIKEN full-length
 enriched library, clone:931018K24 product:solute carrier family 5
 (choline transporter), member 7, full insert sequence.
 ACCESSION AK034415 GI:26329926
 VERSION AK034415.1
 KEYWORDS HTC, CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Mus.
 REFERENCE 1 Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 PUBMED 10349636
 REFERENCE
 AUTHORS 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
 PUBMED Genome Res. 10 (10), 1617-1630 (2000)
 REFERENCE 11042159
 AUTHORS 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system -384-format
 JOURNAL sequencing pipeline with 384 multichannel sequencer
 PUBMED Genome Res. 10 (11), 1757-1771 (2000)
 REFERENCE 11076861
 AUTHORS 4
 TITLE The RIKEN Genome Exploration Research Group Phase II Team and the
 JOURNAL FANTOM Consortium.
 REFERENCE Functional annotation of a full-length mouse cDNA collection
 PUBMED Nature 409, 685-690 (2001)
 REFERENCE
 AUTHORS 5
 TITLE The PANTOM Consortium, the RIKEN Genome Exploration Research Group
 JOURNAL Phase I and II Team.
 REFERENCE Analysis of the mouse transcriptome based on functional annotation
 PUBMED of 60,770 full-length cDNAs
 REFERENCE Nature 420, 563-573 (2002)
 REFERENCE
 AUTHORS 6
 TITLE RIKEN Genome Exploration Research Group, Genome Science Group
 JOURNAL (Genome Network Core Team) and the FANTOM Consortium.
 REFERENCE Antisense transcription in the Mammalian Transcriptome
 PUBMED Science 309, 1564-1566 (2005)
 REFERENCE
 AUTHORS 7
 TITLE The PANTOM Consortium, Riken Genome Exploration Research Group and
 JOURNAL Genome Science Group (Genome Network Project Core Group).
 REFERENCE The Transcriptional Landscape of the Mammalian Genome
 PUBMED Science 309, 1559-1563 (2005)
 REFERENCE
 AUTHORS 8 (bases 1 to 4306)
 TITLE Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, D., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakatsume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M., and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Sueno-cho, Tsurumi-ku, Yokohama,
 Kanagawa, 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
 URL: http://genome-gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome-gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/
 Location/Qualifiers

```
source
1. 4306
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9330188K24"
/db_xref="taxon:10090"
/clone="9330188K24"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
394_2136
/note="unnamed protein product; putative
solute carrier family 5 (choline transporter), member 7
(MGDI[MGI:1927126 GB|Nm_022025, evidence: BLASTN, 99%,
match=1743])"
/codon_start=1
/protein_id="BAC28702.1"
/db_xref="GI:26329927"
/translation="MSFVHESGLVAIIILFYLLILFVGIWAMKTKNSGNPERSEAIIV
GGRDIGLVGGFTMTAVWGGYINGTAEAVVPGCGLAMAQAPIGYSLILGLIF
AKPMRSKYVTMLDPFOOYGRMGGLFI PALMGEEMFAAIFALGATTSVIDV
NVISIVSALLAIIYTLVGLYSVAITDVOLFCIFIGLMTISVPALSHPATVDIGFT
AVHAKYOSPMWLTIESVEVYTMIDNFIIMGSIPIWOAYFORVSSSATYATDSFL
AAFGCTVMAIPAICGAIGASTDMNOTAYGFPDKTKEADNIIIVLYOICPVYISF
FGLGAVSAVMSADSSILSASWFAKNYQUSFRONASDKIIVWMTITVAVFASAT
TAMALLTKTVGLMYLSSDLYIIFPOLCVLFIKGNITVAGVAYIFGLFRLITGG
EPYLILOPLIFPGYSDKNGLYINRFPEKTLMTVSTFNICVSYLAKLYFESGTLF
PKLDVEAVAHSEEMDKITILVRENMIKINELAPVKPRQSLTISFTNKEALLDV
DSPSGSGEDNDLQ"
```

Query Match 99.4%; Score 1731.8; DB 6; Length 4306;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1736; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
1 ATGCTTTTCCATGTGGAAGAGCTGTGAGTATTTATCTTCTTCTTCTTATTTTCG 60
Db ATGCTTTTCCAGTGAAGAGCTGTGAGTATTTATCTTCTTCTTATTTTCG 453
QY GTTGAATATGAGGCTGCATGAAAAACAAAGAGGCAAGCCAGAGAGGCAAGTGA 120
Db GTTGAATATGAGGCTGCATGAAAAACAAAGAGGCAAGCCAGAGAGGCAAGTGA 513
QY GTTGAATATGAGGCTGCATGAAAAACAAAGAGGCAAGCCAGAGAGGCAAGTGA 180
Db GTTGAATATGAGGCTGCATGAAAAACAAAGAGGCAAGCCAGAGAGGCAAGTGA 573
QY ACCGTGGTTGGAGAGGCTACATCAATGGGACAGAGCACTGATAGGCGCAGTTCT 240
Db ACCGTGGTTGGAGAGGCTACATCAATGGGACAGAGCACTGATAGGCGCAGTTCT 633
QY GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTCTG 300
Db GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTCTG 693
QY GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTCTG 360
Db GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTCTG 753
QY TTTTTCGCAAACTATGCGTTCCAAAGGATATGTGATGTATGATAGCCATTTTCAACAG 420
Db TTTTTCGCAAACTATGCGTTCCAAAGGATATGTGATGTATGATAGCCATTTTCAACAG 813
QY ATCTATGAAAGCGATGGGTGGGCTGCTTATCCCTGCACTGATGGGAGAGATGTC 480
Db ATCTATGAAAGCGATGGGTGGGCTGCTTATCCCTGCACTGATGGGAGAGATGTC 873
QY TGGGCTGAGCAATTTTCTGCAATTAAGGGCCACCATCAGCGTATCTATGATGAT 540
Db TGGGCTGAGCAATTTTCTGCAATTAAGGGCCACCATCAGCGTATCTATGATGAT 933
QY GTGAACATATCGGTATGCTCTGCACTCATTTGATCCCTAGTGGGTGGG 600
Db GTGAACATATCGGTATGCTCTGCACTCATTTTATAGACCTGTGG 600
```

```
Db CTCTACTGTGGCATATATGATGTGTCCAGCTATTTCTGCAATTTTATAGACCTGTGG 600
QY 934 CTCTACTGTGGCATATATGATGTGTCCAGCTATTTCTGCAATTTTATAGACCTGTGG 993
QY ATCAGTGTCCCTTTTGGCCCTGTCAATCTCTGACAGTCAAGCCAGATCGATTCGAGCTGTG 660
Db ATCAGTGTCCCTTTTGGCCCTGTCAATCTCTGACAGTCAAGCCAGATTCGAGCTGTG 1053
QY CATGCTAAATATACAGAGTCCCTGGTGGGAACCATTTGAATCATGTTGAAGTCTACACCTGG 720
Db CATGCTAAATATACAGAGTCCCTGGTGGGAACCATTTGAATCATGTTGAAGTCTACACCTGG 1113
QY CTGTGATATTTTCTGTATTTGATGCTGGGTGGAATCCCATGAGCAAGCTACTTCAGAGG 780
Db CTGTGATATTTTCTGTATTTGATGCTGGGTGGAATCCCATGAGCAAGCTACTTCAGAGG 1173
QY GTCTCTCTTCAATCTCTGACCAACCTATGCTGAGTACTGTCTTCTCTGACGCTTTTGGG 840
Db GTCTCTCTTCAATCTCTGACCAACCTATGCTGAGTACTGTCTTCTCTGACGCTTTTGGG 1233
QY TCCCTGTGATGCTCTACCCGCAATATGCAATGAGCTATTTGAGCTTCCACAGACTGG 900
Db TCCCTGTGATGCTCTACCCGCAATATGCAATGAGCTATTTGAGCTTCCACAGACTGG 1293
QY TCCCTGTGATGCTCTACCCGCAATATGCAATGAGCTATTTGAGCTTCCACAGACTGG 960
Db TCCCTGTGATGCTCTACCCGCAATATGCAATGAGCTATTTGAGCTTCCACAGACTGG 1353
QY CCGATGCTTGCAGTACCTGCTGCTGCTGATACATCTCTTCTTGGGCTTGTGCTGTT 1020
Db CCGATGCTTGCAGTACCTGCTGCTGCTGATACATCTCTTCTTGGGCTTGTGCTGTT 1413
QY TCAAGTGTGATGCTCTACCTGCTGCTGATACATCTCTTCTTGGGCTTGTGCTGTT 1080
Db TCAAGTGTGATGCTCTACCTGCTGCTGATACATCTCTTCTTGGGCTTGTGCTGTT 1473
QY TCAAGTGTGATGCTCTACCTGCTGCTGATACATCTCTTCTTGGGCTTGTGCTGTT 1473
Db TCAAGTGTGATGCTCTACCTGCTGCTGATACATCTCTTCTTGGGCTTGTGCTGTT 1533
QY CCGAATATCTACCAAGCTTCTCTGACAAATATGATACAGCAAGAAATGTGAGGCTC 1140
Db CCGAATATCTACCAAGCTTCTCTGACAAATATGATACAGCAAGAAATGTGAGGCTC 1533
QY ATGAGATATCATCTGCTTGTGTCCAGACATCTGCAAGAGCCATGCTTGTGCTGAGAG 1200
Db ATGAGATATCATCTGCTTGTGTCCAGACATCTGCAAGAGCCATGCTTGTGCTGAGAG 1593
QY ACTGTGTATGAGGCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1260
Db ACTGTGTATGAGGCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1653
QY CTGCTCTGTGATCTTTCATCAAGAGAACCAACACTTATGAGGAGCTGTGCTGATATTT 1320
Db CTGCTCTGTGATCTTTCATCAAGAGAACCAACACTTATGAGGAGCTGTGCTGATATTT 1713
QY TTTGAGATATTTCTGAGATTAATGAGAGAGAGCCATATCTATCTGACGCCCTTAAATC 1380
Db TTTGAGATATTTCTGAGATTAATGAGAGAGAGCCATATCTATCTGACGCCCTTAAATC 1773
QY TTTTACCTCGTGTATTAATCTGACAAAGATGTATATACATCAGAGGTTCCCATTTTAA 1440
Db TTTTACCTCGTGTATTAATCTGACAAAGATGTATATACATCAGAGGTTCCCATTTTAA 1833
QY TTTTACCTCGTGTATTAATCTGACAAAGATGTATATACATCAGAGGTTCCCATTTTAA 1500
Db TTTTACCTCGTGTATTAATCTGACAAAGATGTATATACATCAGAGGTTCCCATTTTAA 1893
QY CTATTTGAAAGTGAACCTTGGCTTCAAAATTTAGATGATGATGATGATGATGATGATG 1560
Db CTATTTGAAAGTGAACCTTGGCTTCAAAATTTAGATGATGATGATGATGATGATGATG 1953
QY CACAGTGAAGAGAAATGAGCAAGACATTTCTAGTCAAGAAATGAATTAATTAAT 1620
Db CACAGTGAAGAGAAATGAGCAAGACATTTCTAGTCAAGAAATGAATTAATTAAT 2013
QY GAACTTGACCTGTGAAACCTTGGCAGAGCTTAACTTCACTTCACTTCACTTCACTTAAAG 1680
Db GAACTTGACCTGTGAAACCTTGGCAGAGCTTAACTTCACTTCACTTCACTTCACTTAAAG 1680
```

Db 2014 GAACCTGCACCTGTGAAACCTCGGACAGCCTAACCTCAGTTCAACTTTCACCAATAAG 2073
QY 1681 GAGGCCCTCGTGTATGATTCAGTCCGAGGAGGCTGGGACTGAAGTAATTTACAA 1740
Db 2074 GAGGCCCTCGTGTATGATTCAGTCCGAGGAGGCTGGGACTGAAGTAATTTACAA 2133
QY 1741 TGA 1743
Db 2134 TGA 2136

RESULT 3
AY413300 1743 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM4844 gene, VIRUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY413300
VERSION AY413300.1 GI:39769262
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1743)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
REFERENCE 2 (bases 1 to 1743)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..1743
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1743
/locus_tag="HCM4844"

ORIGIN
gene
Query Match 84.1%; Score 1465; DB 14; Length 1743;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1468; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 1 ATGCCCTTCATGGAAGACCTGTAGCTATATCCTCTTCACTCCCTTATATTCG 60
Db 1 ATGCTCTTCCACGTAGAAGACTGTGTATATCCTCTTCACTCCCTTATATTCG 60
QY 61 GTTGGAATATGGGCTGATGGAAGAACCAAAACAGGCGCAACCAAGAGCGCAGTAA 120
Db 61 GTTGGAATATGGGCTGATGGAAGAACCAAAACAGGCGCAACCAAGAGCGCAGTAA 120
QY 121 GCATATATAGTGGGGGCGCTGACATTTGTTGTTGTTGTTTACATGACAGCC 180
Db 121 GCATATATAGTGGGGGCGCTGACATTTGTTGTTGTTTACATGACAGCC 180
QY 121 GCATATATAGTGGGGGCGCTGACATTTGTTGTTGTTTACATGACAGCC 180
Db 121 GCATATATAGTGGGGGCGCTGACATTTGTTGTTGTTTACATGACAGCC 180
QY 181 ACCTGGGTTGGAAGGCTACATCAATGTGACAGACAGAGTGTATGGGCGGTTGT 240
Db 181 NNN 240
QY 241 GGTCTAGCTTTGGGCTCATGACCCATGTGATTTCTCTGAGCTAATTTTAGGTGCTG 300

Db 241 NNN 300
QY 301 TTTTTCGCAACCTATGCGTTCACAGGATATGTGACTATGTTAGACCATTCACAAACG 360
Db 301 NNN 360
QY 361 ATCTATGAAAGCGCATGGGCTGCTCTTATCCTGCACTGATGGAGAGATGTTTC 420
Db 361 NNN 420
QY 421 TGGGCTGACGCAATTTCTCTGATAGGAGGCGCACCATCATGCTGATGATGAT 480
Db 421 NNN 480
QY 481 GTGAACATATCGGTATGTTCTGCACTCATTTGCACTTTATATCCCTAGGGGTTGG 540
Db 481 GTGAACATATCGGTATGTTCTGCACTCATTTGCACTTTATATCCCTAGGGGTTGG 540
QY 541 CTCTACTCTGGCATATACGTATGTTGTCAGCTATTCGACTTTTATAGGACTGTGG 600
Db 541 CTCTACTCTGGCATATACGTATGTTGTCAGCTATTCGACTTTTATAGGACTGTGG 600
QY 601 ATCAGTCTCTTTGCTGCTGCACTCTGCACTGCACTGCACTGCACTGCTG 660
Db 601 ATCAGTCTCTTTGCTGCTGCACTCTGCACTGCACTGCACTGCACTGCTG 660
QY 661 CATGCTAATATCCAGAGTCCCTGCTGCGGAGCAACATGAACTGAAAGCTACCTGG 720
Db 661 CATGCTAATATCCAGAGTCCCTGCTGCGGAGCAACATGAACTGAAAGCTACCTGG 720
QY 721 CTGTATATTTCTGTATGATGCTGGGTGAAATCCCATGCGAAGCTTCTTCAAGG 780
Db 721 CTGTATATTTCTGTATGATGCTGGGTGAAATCCCATGCGAAGCTTCTTCAAGG 780
QY 781 GTCTCTCTTCATCTGACGACCATGATGCTGAGTATGCTCTTCTGCGAGCTTTGG 840
Db 781 GTCTCTCTTCATCTGACGACCATGATGCTGAGTATGCTCTTCTGCGAGCTTTGG 840
QY 841 TGGCTGTGATGGCTGTACCCGCAATGATGATGAGGCTTATGAGCTTCAAGACTGG 900
Db 841 TGGCTGTGATGGCTGTACCCGCAATGATGATGAGGCTTATGAGCTTCAAGACTGG 900
QY 901 AACCAAGTCCCTACGAGGATCCAGATCCCAAGACTTAAAGAGGAGCAGATGATTC 960
Db 901 AACCAAGTCCCTACGAGGATCCAGATCCCAAGACTTAAAGAGGAGCAGATGATTC 960
QY 961 CCGATGCTTCTGAGTACCTTGGCTGTGTATCATCTCTTCTTGGGCTTGGTCTGT 1020
Db 961 CCGATGCTTCTGAGTACCTTGGCTGTGTATCATCTCTTCTTGGGCTTGGTCTGT 1020
QY 1021 TCACTGCTGTATGCTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 TCACTGCTGTATGCTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 CCGAATATCTACAGCTTCTTCAAGAAATGATCAGACAAAGAAATTTGTGGGCTC 1140
Db 1081 CCGAATATCTACAGCTTCTTCAAGAAATGATCAGACAAAGAAATTTGTGGGCTC 1140
QY 1141 ATGAGGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1141 ATGAGGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1201 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db 1201 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
QY 1261 CTGCTCTGTGACTCTTCAACAAAGAACCAACACTTATGAGGAGCACTGCTGCTAT 1320
Db 1261 CTGCTCTGTGACTCTTCAACAAAGAACCAACACTTATGAGGAGCACTGCTGCTAT 1320
QY 1321 TTTGACTATCTCTGAGATTAATGAGAGAGACCATATCTAATCTGAGCCCTTAATC 1380

Db	1321	TTTGACATATTCTAGAAATTA	CTGAGAGAGACCATTCATTA	CTTCAGACCCCTAATC	1380
Qy	1381	TTTCAACCTGGTATTA	CTTGACAGAAATGGTATATAC	ATACAGAGTTCCTCATTTAA	1440
Db	1381	TTTCAACCTGGTATTA	CTTGACAGAAATGGTATATAC	AAACAGAGGTCCCATTTAA	1440
Qy	1441	ACTCTCCATAGTATACCTCA	TTCTTTACCAACATTTTGTTCTT	ATCTACACCAAGTAT	1500
Db	1441	ACTCTCCATAGTATACCTCA	TTCTTTACCAACATTTTGTTCTT	ATCTACACCAAGTAT	1500
Qy	1501	CTATTGAAAGTGAACCTTG	CTCCCAAAATTAGATGATTTG	ATGATGCTGTTGTGCAAGG	1560
Db	1501	CTATTGAAAGTGAACCTTG	CTCCCAAAATTAGATGATTTG	ATGATGCTGTTGTGCAAGG	1560
Qy	1561	CACAGTGAAGAGAACATG	AGACACCACTTCAGCAGAAATG	AAATATCAAAATTAAT	1620
Db	1561	CACAGTGAAGAGAACATG	AGACACCACTTCAGCAGAAATG	AAATATCAAAATTAAT	1620
Qy	1621	GAACCTGCACCTGTGAAA	CTTGCGAGAGCCCTAACCTC	AGTTCAACTTCACCAATAAG	1680
Db	1621	GAACCTGCACCTGTGAAA	CTTGCGAGAGCCCTAACCTC	AGTTCAACTTCACCAATAAG	1680
Qy	1681	GAGGCCCTCTTGATGTTG	ATTCACGTCGCGAGGGGTCT	GGGACCTGAAGATTAACA	1740
Db	1681	GAGGCCCTCTTGATGTTG	ATTCACGTCGCGAGGGGTCT	GGGACCTGAAGATTAACA	1740
Qy	1741	TGA	1743		
Db	1741	TGA	1743		
RESULT 4	AY413298	1743 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY413298		Homo sapiens HCM4844 gene, VIRTUAL	TRANSCRIPT, partial	sequence,
DEFINITION					
ACCESSION	AY413298				
VERSION	AY413298.1	GI:39769260			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
REFERENCE	1 (bases 1 to 1743)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,				
	Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,				
	Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,				
	Adams,M.D. and Cargill,M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous				
	gene trilos				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 1743)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,				
	Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,				
	Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,				
	Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,				
	Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering				
	them based on alignment				
FEATURES	Location/Qualifiers				
source	1..1743				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
gene	/db_xref="taxon:9606"				
	<1..>1743				
	/locus_tag="HCM4844"				
ORIGIN					
Query Match	66.7%;	Score 1162.6;	DB 14;	Length 1743;	

Qy	1	ATGCTTTCCATGTGGAAGAG	CTGGTAGCTATTAATCTCTT	CTTACCTCTTAATTTCTG	60
Db	1	ATGCTTTCCATGTGGAAGAG	CTGGTAGCTATTAATCTCTT	CTTACCTCTTAATTTTCTG	60
Qy	61	GTGGAATATGGGCTGATG	GAAGAAACAAAAACAGCGG	CAACCAAGAGCGCAGTAA	120
Db	61	GTGGAATATGGGCTGATG	GAAGAAACAAAAACAGTGG	CAACCAAGAGCGCAGGAA	120
Qy	121	GCCATCATAGTGGGTGCG	CTGGAGACCAAAAACGTGG	CAACCAAGAGCGCAGGAA	180
Db	121	GCCATCATAGTGGGTGCG	CTGGAGACCAAAAACGTGG	CAACCAAGAGCGCAGGAA	180
Qy	181	ACCTGGTTGGAGAGGCTA	CAATCATGGGACAGAACAG	ACAGTATGGGCCAGTGT	240
Db	181	ACCTGGTTGGAGAGGCTA	CAATCATGGGACAGAACAG	ACAGTATGGGCCAGTGT	240
Qy	241	GGTCTACCTTGGGCTCAT	GACCCATGAGATATCTCT	GTGAGTAAATTTAGTGCTG	300
Db	241	GGTCTACCTTGGGCTCAT	GACCCATGAGATATCTCT	GTGAGTAAATTTAGTGCTG	300
Qy	301	TTTTTTCGAAACCTATG	CGCTTCCAGGAGATATGTG	ACTATGTTAGACCAATTC	360
Db	301	TTTTTTCGAAACCTATG	CGCTTCCAGGAGATATGTG	ACTATGTTAGACCAATTC	360
Qy	361	ATCTATGAAAGCGCAT	GTGGGTGGGCTCTCTT	CATCCCTGACATGAGAG	420
Db	361	ATCTATGAAAGCGCAT	GTGGGTGGGCTCTCTT	CATCCCTGACATGAGAG	420
Qy	421	TGGCTCGACGAATTTCT	CTGCAATTAGGGCCACCAT	CACGCTGATCATTTATG	480
Db	421	TGGCTCGACGAATTTCT	CTGCAATTAGGGCCACCAT	CACGCTGATCATTTATG	480
Qy	481	GTGAACATATCGGCAT	TTGTCTCGACCTGATGGCA	TTCTTTATACCTAGTGG	540
Db	481	GTGAACATATCGGCAT	TTGTCTCGACCTGATGGCA	TTCTTTATACCTAGTGG	540
Qy	541	CTTACTCTGTGCATAT	CTGATGTTGTGCAGCTAT	TTCGATTTTATAGACTGTG	600
Db	541	CTTACTCTGTGCATAT	CTGATGTTGTGCAGCTAT	TTCGATTTTATAGACTGTG	600
Qy	601	ATCAGTGTCCCTTTTGG	CCCTGTGCATCTCTGCACT	CAACGATCGGATTCACAG	660
Db	601	ATCAGTGTCCCTTTTGG	CCCTGTGTGCATCTCTGCA	CTCCAGCTGCAATCGGCTG	660
Qy	661	CATGCTAAATACCGAGAG	CCCTGGCTGGGAACCTGAT	GAATCAAGTGAAGTCAAC	720
Db	661	CATGCTAAATACCGAGAG	CCCTGGCTGGGAACCTGAT	GAATCAAGTGAAGTCAAC	720
Qy	721	CTTGATTAATTTTCTG	TATTGATGCTGGTGGAA	TCCCATGAGCAAGCTTAC	780
Db	721	CTTGATTAATTTTCTG	TATTGATGCTGGTGGAA	TCCCATGAGCAAGCTTAC	780
Qy	781	GTCCTCTTCATGCTAG	CAACCTATGTCACAGTAC	TGTCCTTCCTGGAGCTTTG	840
Db	781	GTCCTCTTCATGCTAG	CAACCTATGTCACAGTAC	TGTCCTTCCTGGAGCTTTG	840
Qy	841	TGCTGTGATGAGCTTAC	CCGACCATATGATAGAACT	ATGAGACTTATGAGGCTT	900
Db	841	TGCTGTGATGAGCTTAC	CCGACCATATGATAGAACT	ATGAGACTTATGAGGCTT	900
Qy	901	AACCAAGACTGCGCTA	CGGGTATCCAGATCCCAAG	ACTTAAGAGAGAAAGCAT	960
Db	901	AACCAAGACTGCGCTA	CGGGTATCCAGATCCCAAG	ACTTAAGAGAGAAAGCAT	960
Qy	961	CCGATGCTTGTGAGTAC	TCTGTGCGCTGTGTACAT	CTCTTTTGGGCTTGGTCT	1020
Db	961	CCATTTGTTGTGAGTAC	TCTGTGCGCTGTGTATAT	TTCTTTTGGTCTTGGTCA	1020
Qy	1021	TGAGCTGTGCATGTCCT	GACTGACTGTGCTCATCT	GTCGTGGCGAGTCTTAT	1080
Db	1021	TGAGCTGTGCATGTCCT	GACTGACTGTGCTCATCT	GTCGTGGCGAGTCTTAT	1080

Db 1021 TCTGCTGCTGATATGATCATCAGACAGATTCCTTCATCTTGTCAGCAAGTTCATGTTTGA 1080
 Oy 1081 CGGAATATCATCAGACCTTCTCTTCCAGCAAAATGATCAGACAGAGAAATGTTGGGTC 1140
 Db 1081 CGGAATATCATCAGACCTTCTCTTCCAGCAAAATGATCAGACAGAGAAATGTTGGGTC 1140
 Oy 1141 ATGAGATCATCTGTGCTGTGTGTCGAGCATCTGCAACAGCCATGCTTGTCTGACGAAG 1200
 Db 1141 ATGAGATCATCTGTGCTGTGTGTCGAGCATCTGCAACAGCCATGCTTGTCTGACGAAG 1200
 Oy 1201 ACTGTGTATGAGGCTCTGTGTAAGCTGAGCTGACCTTGTCTACTCATCATCTTCCACAG 1260
 Db 1201 ACTGTGTATGAGGCTCTGTGTAAGCTGAGCTGACCTTGTCTACTCATCATCTTCCACAG 1260
 Oy 1261 CTGCTCTGTGTAAGCTGTAAGGAGACCAACATTAATGAGGCTGTGCTGTATAT 1320
 Db 1261 CTGCTCTGTGTAAGCTGTAAGGAGACCAACATTAATGAGGCTGTGCTGTATAT 1320
 Oy 1321 TTTGACATATTCCTGAGAAATTAAGTGAAGAGAGACCATATCTATATCTTGACGCTTAAATC 1380
 Db 1321 TTTGACATATTCCTGAGAAATTAAGTGAAGAGAGACCATATCTATATCTTGACGCTTAAATC 1380
 Oy 1381 TTTGACATATTCCTGAGAAATTAAGTGAAGAGAGACCATATCTATATCTTGACGCTTAAATC 1440
 Db 1381 TTTGACATATTCCTGAGAAATTAAGTGAAGAGAGACCATATCTATATCTTGACGCTTAAATC 1440
 Oy 1441 ACTCTCTGATGCTGTAAGCTGTAAGGAGACCAACATTAATGAGGCTGTGCTGTATAT 1500
 Db 1441 ACTCTCTGATGCTGTAAGCTGTAAGGAGACCAACATTAATGAGGCTGTGCTGTATAT 1500
 Oy 1501 CTATTTGAAAGTGAAGCTTGTCTGCAAAATTAAGTGAAGAGAGACCATATCTTGACGCTTAAATC 1560
 Db 1501 CTATTTGAAAGTGAAGCTTGTCTGCAAAATTAAGTGAAGAGAGACCATATCTTGACGCTTAAATC 1560
 Oy 1561 CACAGTGAAGAGAACTGCAAGACCATTTCTGTCAGAAATGAATTAATTAATTAAT 1620
 Db 1561 CACAGTGAAGAGAACTGCAAGACCATTTCTGTCAGAAATGAATTAATTAATTAAT 1620
 Oy 1621 GAACCTGCACTGTAAGCTGTAAGGAGACCAACATTAATGAGGCTGTGCTGTATAT 1680
 Db 1621 GAACCTGCACTGTAAGCTGTAAGGAGACCAACATTAATGAGGCTGTGCTGTATAT 1680
 Oy 1681 GAGGCTCTCTGATGATGATTCAGTCCGAGAGGCTGTGAGCTAAGATTAATTAAT 1740
 Db 1681 GAGGCTCTCTGATGATGATTCAGTCCGAGAGGCTGTGAGCTAAGATTAATTAAT 1740
 Oy 1741 TGA 1743
 Db 1741 TGA 1743

RESULT 5
 AY413299 1743 bp DNA linear GSS 12-DEC-2003
 LOCUS Pan troglodytes HCM444 gene, VIRUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY413299
 VERSION AY413299.1 GI:39769261
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 BUKARYOTA: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 MAMMALIA: Eutheria; Euarchontoglires; Primates; Catarrhini;
 HOMINIDAE; Pan. (bases 1 to 1743)
 REFERENCE 1 (bases 1 to 1743)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tannenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302

REFERENCE 2 (bases 1 to 1743)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tannenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment
 FEATURES
 source
 1..1743
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>1743
 /locus_tag="HCM4844"
 ORIGIN
 Query Match 66.7%; Score 1162.2; DB 14; Length 1743;
 Best Local Similarity 73.3%; Pred. No. 0;
 Matches 1278; Conservative 0; Mismatches 465; Indels 0; Gaps 0;
 Oy 1 ATGCCCTTTCATGTGAGAGACTGTGATATATCTCTTCTTACCTCTTATATTTCTG 60
 Db 1 ATGCCCTTTCATGTGAGAGACTGTGATATATCTCTTCTTACCTCTTATATTTCTG 60
 Oy 61 GTTGGAAATATGGGCTGCATGGAAGAAACCAAAACAGGGGCAACCCAAAGAGCGCAGTGA 120
 Db 61 GTTGGAAATATGGGCTGCATGGAAGAAACCAAAACAGGGGCAACCCAAAGAGCGCAGTGA 120
 Oy 121 GCCATCATAGTGTGGGCGGTGACATTTGTTGGTGTGTTTACATGACAGCC 180
 Db 121 GCCATCATAGTGTGGGCGGTGACATTTGTTGGTGTGTTTACATGACAGANN 180
 Oy 181 ACTGTGTTGAGAGAGCTATCATGTGGAACAGCAAGACAGTGTATGGGCGAGTTGT 240
 Db 181 NNN 240
 Oy 241 GGCTAGCTGGGCTCATGCAACCATGTGATATCTCTAGTCTAATTTAGTGGTCTG 300
 Db 241 NNN 300
 Oy 301 TTTTTCGCAACCTATGCGTTCAGAGATATGTGACTATGTAGACCATCAAAACAG 360
 Db 301 NNN 360
 Oy 361 ATCTATGAAAGCGCATGGGTGGGCTCTTCTATCCCTGCACTGATGGAGAGATGTTT 420
 Db 361 NNN 420
 Oy 421 TGGGCTGCAGCAATTTCTCTGCAATTAAGGGCCACATCAGCGTATCATATGTGGAT 480
 Db 421 NNN 480
 Oy 481 GTGAACATATCGGTGATGTCCTGCACTATGACATCTTATATACCTAGTGGTGG 540
 Db 481 GTGAACATATCGGTGATGTCCTGCACTATGACATCTTATATACCTAGTGGTGG 540
 Oy 481 ATGCACATTTCTGTATCATCTCTGCACTATGACATCTGTCACACTGTGAGAGG 600
 Db 481 ATGCACATTTCTGTATCATCTCTGCACTATGACATCTGTCACACTGTGAGAGG 600
 Oy 541 CTCTACTCTGTGGCATATATGATGTTGTCAGCTATCTGCACTTTTATAGAGCTGTG 600
 Db 541 CTCTATTTCTGTGGCCTTACACTGATGTGTTCAAGCTCTTTTGAATTTTGAAGGCTGTG 600
 Oy 601 ATCAGTGTCCCTTGTGATGTCATCTGCACTATGACATCTGTCACACTGTGAGAGG 660
 Db 601 ATCAGTGTCCCTTGTGATGTCATCTGCACTATGACATCTGTCACACTGTGAGAGG 660
 Oy 661 CATGCTAAATACAGAGTCCCTGCTGGAGCAACATGATGATGATGATGATGATGATGAT 720
 Db 661 CATGCTAAATACAGAGTCCCTGCTGGAGCAACATGATGATGATGATGATGATGATGAT 720
 Oy 721 CTGTGATATTTTCTGTATGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAG 780
 Db 721 CTGTGATATTTTCTGTATGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAG 780

Db 721 CTTGATGATTTTCTGTTGTTGATGCTGGGTGGAATCCATGGCAAGCATCTTTCAGAG 780
 Oy 781 GTCTCTCTTCACTCTCAGCCACCTATGCTCAGTACTGCTCTCTGACAGCTTTTGGG 840
 Db 781 GTTCTCTCTTCTCTCAGCCACCTATGCTCAGTACTGCTCTCTGACAGCTTTTGGG 840
 Oy 841 TGCCCTGCTGATGCTCTACCCGCCATATGACATAGAGAGCTTATGGAGCTTCCACAGACTGG 900
 Db 841 TGCCCTGCTGATGCTCTACCCGCCATATGACATAGAGAGAGCTTATGGAGAGCTTCCACAGACTGG 900
 Oy 901 AACAGACTGCTCAGGGTATCCAGATCCCAAGACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 Db 901 AACAGACTGCTCAGGGTATCCAGATCCCAAGACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 Oy 961 CCGATGCTTCTGACAGTACTCTGACCTGCTGTAATCTCTCTTGGGCTTGGTCTGTT 1020
 Db 961 CCGATGCTTCTGACAGTACTCTGACCTGCTGTAATCTCTCTTGGGCTTGGTCTGTT 1020
 Oy 1021 TCAGTCTGCTGATGCTCTCAGCTGACTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 1021 TCAGTCTGCTGATGCTCTCAGCTGACTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Oy 1081 CGGAATATCTACACAGCTTCTCTGAGACAAATGATCAGACAGAGAAATGCTGGGCTC 1140
 Db 1081 CGGAATATCTACACAGCTTCTCTGAGACAAATGATCAGACAGAGAAATGCTGGGCTC 1140
 Oy 1141 ATGAGAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Db 1141 ATGAGAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Oy 1201 ACTGCT 1260
 Db 1201 ACTGCT 1260
 Oy 1261 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 Db 1261 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 Oy 1321 TTTGAGATTAATCTGAGAAATTAAG 1380
 Db 1321 TTTGAGATTAATCTGAGAAATTAAG 1380
 Oy 1381 TTTGAGATTAATCTGAGAAATTAAG 1440
 Db 1381 TTTGAGATTAATCTGAGAAATTAAG 1440
 Oy 1441 ACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Db 1441 ACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Oy 1501 CTATTTGAAAGTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 Db 1501 CTATTTGAAAGTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 Oy 1561 CACAGTGAAG 1620
 Db 1561 CACAGTGAAG 1620
 Oy 1621 GAACCTGCACTGCTGCAAG 1680
 Db 1621 GAACCTGCACTGCTGCAAG 1680
 Oy 1681 GAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 Db 1681 GAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 Oy 1741 TGA 1743
 Db 1741 TGA 1743

RESULT 6
 AKI64116

LOCUS AKI64116 4122 bp mRNA linear HTC 21-SEP-2005
 DEFINITION Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length enriched library, clone: C530033E06 product: isolate carrier family 5 (choline transporter), member 7, full insert sequence.
 VERSION AKI64116
 KEYWORDS AKI64116.1 GI:74211080
 SOURCE HTC, CAP trapper.
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 REFERENCE 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 10349636
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 11042159
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, N., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multiplexed sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 11076861
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Kono, H., Adachi, T., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuwaji, P., Lewis, S., Matsuda, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staub, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. P., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombere, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyokawa, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kotsuki, S., and Hayashizaki, Y.
 RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 11217851

5 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Oshino, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chochia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Gilman, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kana, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,

CONSRM	JOURNAL	REFERENCE	AUTHORS
12466851	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Nature 420 (6915), 563-573 (2002)	Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, U.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempke, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Welle, C., Wilming, L.G., Wyszewski, B., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayashizaki, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
CONSRM	JOURNAL	REFERENCE	AUTHORS
12466851	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Nature 420 (6915), 563-573 (2002)	Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Imbabiato, A., Arner, J., Attaluri, R., Auer, R., Bailey, T.L., Banerji, M., Baxter, L., Belsel, K.W., Bersano, T., Bonic, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christofidellis, A., Cluttenbuck, D.R., Crome, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., Di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Flecher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgi-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Habers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwano, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kello, J., Kitamura, H., Kitano, H., Kojima, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liut, L.S., McMillan, S., Madan, Babu, M., Madera, M., Marchionni, L., Matuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Moritz, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sanderlin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Sempke, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinalati, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yang, K., Yamashita, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Bruce, V., Quackenbush, J., Wahlestedt, C., Watters, D., Wells, D.A., Kall, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Aizawa, T., Iida, J., Imanura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Kono, H., Nakano, K., Nimmiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanabe, K., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.
CONSRM	JOURNAL	REFERENCE	AUTHORS
16141072	The transcriptional landscape of the mammalian genome	Science 309 (5740), 1559-1563 (2005)	Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakaniishi, M., Nakamura, M., Nishida, H., Iap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Matlick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Fagiolini, M., Sanderlin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.
CONSRM	JOURNAL	REFERENCE	AUTHORS
1069	RIKEN Genome Exploration Research Group	Antisense transcription in the mammalian transcriptome	

Db 486 TCATGTTGCTCGGAATATCTACCAAGCTTCTTCCAGCAAAATGATCAGACAGGA 545

QY 1129 ATTGTGGGTATGAGGATCATCTGTGCTTGTGTGAGAGATCTGCAACAGCCATGGCT 1188

Db 546 ATGTGTGGGTATGAGGATCATCTGTGCTTGTGTGAGAGATCTGCAACAGCCATGGCT 605

QY 1189 TTGCTGAGGAAGACGTGTATGGGCTCTGTGTAACCTGAGCTCTGACCTTGTCTACATCATC 1248

Db 606 TTGCTGAGGAAGACGTGTATGGGCTCTGTGTAACCTGAGCTCTGACCTTGTCTACATCATC 665

QY 1249 ATCTCCACAGCTGCTGTGTACTCTTCATCAAGGAACCAAGCACTTATGGGAGATT 1308

Db 666 ATCTCCACAGCTGCTGTGTACTCTTCATCAAGGAACCAAGCACTTATGGGAGATT 725

QY 1309 GCTGTTATATTTTGGACTATCTGAGAAATTAATCTGAGAGAGAGCCATATATATCTTG 1368

Db 726 GCTGTTATATTTTGGACTATCTGAGAAATTAATCTGAGAGAGAGCCATATATATCTTG 785

QY 1369 CAGCCCTTAATTTCTTACCTGTTATTAATCTGACAGAAATGTATATCATCAGAG 1428

Db 786 CAGCCCTTAATTTCTTACCTGTTATTAATCTGACAGAAATGTATATCATCAGAG 845

QY 1429 TTCCATTTAAATCTCTCCATGTTACCTTATCTTACCAACATTTGTTCTTAT 1488

Db 846 TTCCATTTAAATCTCTCTCCATGTTACCTTATCTTACCAACATTTGTTCTTAT 905

QY 1489 CTAGCCAGTATCTATTTGAAAGTGAACCTTCCCTCCAAATTAATGATATTTGATGCT 1548

Db 906 CTAGCCAGTATCTATTTGAAAGTGAACCTTCCCTCCAAATTAATGATATTTGATGCT 965

QY 1549 GTTGTGCGAAGGACAGTAAGAAACATGACAGAACATTTCTTACGAAATGAAT 1608

Db 966 GTTGTGCGAAGGACAGTAAGAAACATGACAGAACATTTCTTACGAAATGAAT 1025

QY 1609 ATCAATTAATGAATCTGACCTGTGAAACCTCGGAGAGGCTTAACCTCTGATCACT 1668

Db 1026 ATCAATTAATGAATCTGACCTGTGAAACCTCGGAGAGGCTTAACCTCTGATCACT 1085

QY 1669 TTACCAATAAGAGAGGCTCTCTTGAATGTTGATTCAGTCCGAGAGGCTCTGGAATGAA 1728

Db 1086 TTACCAATAAGAGAGGCTCTCTTGAATGTTGATTCAGTCCGAGAGGCTCTGGAATGAA 1145

QY 1729 GATTAATTAATGAATGA 1743

Db 1146 GATTAATTAATGAATGA 1160

RESULT 7
CD350164 707 bp mRNA linear EST 09-JUL-2003

DEFINITION
UI-M-FY0-cfl-h-10-0-UI-r1 NIH_BMAP_FY0 Mus musculus cDNA clone

ACCESSION
CD350164
IMAGE:6851099 5', mRNA sequence.

VERSION
CD350164.1 GI:31141679

KEYWORDS
EST.

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 707)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabrs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed By: Dr. M. Bento Soares, University of Iowa
DNA Sequencing By: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1..707
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6851099"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="MD10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pyx-Asc; Site: 1: EcoR I;
Site: 2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pyx-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGAGAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Query Match 34.6%; Score 603.4; DB 5; Length 707;
Best Local Similarity 98.4%; Pred. No. 1.8e-163;
Matches 662; Conservative 0; Mismatches 6; Indels 5; Gaps 5;

QY 1 ATGCTTTCCATGTTGGAAGAGACTGTAGCTATATCTCTTCTACCTCTTATATTTCTG 60

Db 27 ATGTCTTCCACGTGAAAGAGCTGTACTATATCTCTTCTACCTCTTATATTTCTG 86

QY 61 GTTGAATATGAGGCTGATGGAACCAAAACAGCGGCAACCGAAGAGCGAGTGA 120

Db 87 GTTGAATATGAGGCTGATGGAACCAAAACAGCGGCAACCGAAGAGCGAGTGA 146

QY 121 GGCATCATAGTCGGGGGCGGTGACATTTGTTGTTGGTGTATACCATGACAGCC 180

Db 147 GGCATCATAGTCGGGGGCGGTGACATTTGTTGTTGGTGTATACCATGACAGCC 206

QY 181 ACCTGGTTGAGAGAGGCTATACATGAGAGAGAGAGAGAGAGAGAGAGAGTGT 240

Db 207 ACCTGGTTGAGAGAGGCTATACATGAGAGAGAGAGAGAGAGAGAGAGAGTGT 266

QY 241 GGTCTAGCTTGGGCTCATGACCCATTGGATATTTCTGAGCTTATATTTAGGTGCTG 300

Db 267 GGTCTAGCTTGGGCTCATGACCCATTGGATATTTCTGAGCTTATATTTAGGTGCTG 326

QY 301 TTTTGTGGAACCAATGCGTTCCAAAGGATATGATGTTAGACCCATTCAAACAG 360

Db 327 TTTTGTGGAACCAATGCGTTCCAAAGGATATGATGTTAGACCCATTCAAACAG 386

QY 361 ATCTATGGAAGCGATGAGGTGGGCTGCTTTCATCTGCACTGATGGG-AGAGATGT 419

Db 387 ATCTATGGAAGCGATGAGGTGGGCTGCTTTCATCTGCACTGATGGGAGAGATGT 446

QY 420 CTGGGCTCAGCAATTTTCTGCAATTAGGGGCGACCA-TCAGCGTATCATT-GATGTG 477

Db 447 CTGGGCTCAGCAATTTTCTGCAATTAGGGGCGACCA-TCAGCGTATCATT-GATGTG 506

QY 478 GATGTGAACATATGCGTATGTCCTGCACTCATTTGCACTTCTTATACCTAGTGGT 537

Db 507 GATGTGAACATATGCGTATGTCCTGCACTCATTTGCACTTCTTATACCTAGTGGT 566

QY 538 GGGCTCTACTCTGTGGCATATAT-GATGTGTCCAGCTATTTGATTTATAGACT 596

Db 567 GGGCTCTACTCTGTGGCATATATGATGTGTGTCCAGCTATTTGATTTATAGACT 626

FEATURES

source

Location/Qualifiers

1. 828
/organism="Sus scrofa"
/mol_type="genomic DNA"
/db_xref="taxon:9623"
/clone="CH242-157K9"
/tissue_type="White blood cells"
/note="Vector pTARBAC1.3_BamHI
sex female"

ORIGIN

Query Match 27.1%; Score 472; DB 14; Length 828;

Best Local Similarity 84.2%; Pred. No. 7.5e-130;

Matches 532; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

1112 ATGATCAGACAGAAATTTGGTGGTATGAGATCACTGTGCTTGTTCGAGACAT
1112 AGGCTTCCGACAGGAGATGCTGTGGTATCGGATCAGATTCGTTGGTGGCT
1112 CTGCAACAGCCATGCTTGTGAGAGAGACGTGTATGGGCTGTGCTGAGCTCG
214 CTGCAACAGCCATGCTTGTGAGAGAGACGTGTATGGGCTGTGCTGAGCTCG
1232 ACCCTGTACATCATCTTCCGACAGCTGCTGTGCTGCTTCTTCAAGAGACCA
274 ACCTGCTACATCATCTTCCGACAGCTGCTGTGCTGCTTCTTCAAGAGACCA
1232 ACCTTATGGGAGTGTGCTGTATTTTGGACTATTTCTGAGATTTACTGAGAG
334 ACACGTACCGGGCCGCGGAGGGATGACATTCGCGCTTTCTGAGGGTAAACGGTGGAG
1332 AGCCATATATCTTTCGACGCCCTTAATCTTACCCGCTTATTAATCTGACAGAAAG
334 AGCCATATCTGAAACCTGACGCCCTTATTTTACCTGCTGTATTAATCTGACAGAAAG
1412 GTATATACATCAGAGATTCCTCATTTAACTCTCCATGCTTACCTTCTTACCA
454 GTATATATATCAGAGATTCCTCATTTAACTCTCCATGCTTACCTTCTTATCCA
1472 ACATTTGCTTTCTTATCTAGCCAAATATCTTAAAGTGGAACTTCTCCAAAT
514 ACATTTGCTTTCTTATCTAGCCAAATATCTTAAAGTGGAACTTCTCCAAAT
1532 TAGATGATTTTGAATGCTGTGTCGACAGGACAGAGAAAGAAATGACCAAGCATTC
574 TAGATATGTTGATGCTGTGTCGACAGGACAGAGAAAGAAATGACCAAGCATTC
1592 TAGTCAAGAAATATCAATTAATGAATGAACTTGTGAACCTCGGACAGAGCC
634 TGCTCAAAATATGAATTAATTAATTAATGAAGAACTTGTGAAGCTCGGACAGAGCC
1632 TAACCTCAGTTCACTTTACCAATTAAGAGAGCCCTCTTGAATGTTGATTCAGTCGG
694 TAACCTCAGTTCACTTTACCAATTAAGAGAGCCCTCTTGAATGTTGATTCAGTCGG
1712 AGGGCTCGGAGCTGAAGTAATTAATTAATGA
754 AAGGTTCTGGGACAGAAATTTGTAATTA
785

RESULT 11

LOCUS

BB626260 RIKEN full-length enriched, adult male diencephalon Mus

musculus cDNA clone 9330170D24 5', mRNA sequence.

BB626260.1 GI:16464298

EST.

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1 (bases 1 to 650)
Arikawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hata,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Kono,M., Konda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takehashi,F.,
Takeeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arikawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Kono,M., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
Wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsunaga,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES

source

Location/Qualifiers

1. 650

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="9330170D24"

/sex="male"

/tissue_type="dienecephalon"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, adult male
dienecephalon"

/note="Site 1: SalI; Site 2: BamHI. cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTTTTCTTTTCTTTTCTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 185.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTATTAATTAATATCCCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pluescript KS(+) after bulk excision
from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:

ORIGIN BamHI"

Query Match 23.5%; Score 409.6; DB 7; Length 650;
Best Local Similarity 96.7%; Pred. No. 4,2e-111;
Matches 437; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 1 ATGCTTTCATGTCGGAAGAGCTGATGATATATCTTCTTACCTCTTATATTTCTG 60
DB 201 ATGCTTTCACGTGGAAGAGCTGATGATATATCTTCTTACCTCTTATATTTCTG 260
QY 61 GTTGAATATGCGGCGTCATGAAACCAAAACACGCGAACCAGGAAGCGCGTGA 120
DB 261 GTTGAATATGCGGCGTCATGAAACCAAAACACGCGAACCAGGAAGCGCGTGA 320
QY 121 GCCATCATATGTCGCGGCGCTGACATGTTGTTGTTGTTGTTTACATGACAGCC 180
DB 321 GCCATCATATGTCGCGGCGCTGACATGTTGTTGTTGTTTACATGACAGCC 380
QY 181 ACTGCGTTGAGAGAGGCTACATCAATGAGACAGAGAGAGTATGCGCCAGTTGT 240
DB 381 ACTGCGTTGAGAGAGGCTACATCAATGAGACAGAGAGAGTATGCGCCAGTTGT 440
QY 241 GGTCTAGCTTGGGCGTCATGACCCATTTGATATCTTGAAGTCTATTTAGTGTCTG 300
DB 441 GGTCTAGCTTGGGCGTCATGACCCATTTGATATCTTGAAGTCTATTTAGTGTCTG 500
QY 301 TTTTTCGGAACCTATGCGGTCGCAAGGATATGATGATGTTAGACCCATTCAACG 360
DB 501 TTTTTCGCG -AACCTATGCGGTCGCAAGGATATGATGATGTTAGACCCATTCAACG 559
QY 361 ATCTATGAAAGCGCATGCGTGGGCTGCTCTTCATCCTGCACTGATGAGAGATGTC 420
DB 560 ATCTATGAAAGCGCATGCGTGGGCTGCTCTTCATCCTGCACTGATGAGAGATGTC 619
QY 421 TGGGCTGACGAATTTCTCTGCATTAGGCGC 452
DB 620 TGGGCTGACGCCA-TTTCTCTGCATTAGGCGC 650

RESULT 12
AG157499 672 bp DNA linear GSS 09-JAN-2002
LOCUS Pan troglodytes DNA, clone: RP43-022H02.T7, genomic survey
DEFINITION
AG157499
AG157499.1 GI:16687177
VERSION
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pan.

REFERENCE
AUTHORS
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Tokoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library RPCI-43
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 672)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Tokoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
COMMENT (E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: T7
LIBRARY

FEATURES
source
Vector 1 : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
Location/Qualifiers
1..672
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-022H02.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 23.3%; Score 406.8; DB 14; Length 672;
Best Local Similarity 84.9%; Pred. No. 3e-110;
Matches 479; Conservative 0; Mismatches 82; Indels 3; Gaps 2;

QY 1110 AATGCAATGACAAAGAAATTTGTTGGGTCATGAGATCACTGTGTTGTTGAGC 1169
DB 111 ACAGGCTTCGACAAAGAAATCGTTGGGTTATGCAATCAAGTGTGTTGTTGAGC 170
QY 1170 ATCTGCAACAGCCATGCGCTTGTGAGAGACTGTGATGAGGCTCTGATCCTTCAACAAAGAAC 1229
DB 171 ATCTGCAACAGCCATGCGCTTGTGAGAGACTGTGATGAGGCTCTGATCCTTCAACAAAGAAC 1289
QY 1230 TGACCTGTCTACATCATCATCTTCCACAGCTGCTGTGTACTCTTCAACAAAGAAC 1289
DB 231 TGACCTGTCTACATCATCATCTTCCACAGCTGCTGTGTACTCTTCAACAAAGAAC 1290
QY 1290 CAACCTATGAGGAGAGTTGCTGTTATATTTTGGACTATTTCCGAGAAATTAAGAGG 1349
DB 291 CAACCTATGAGGAGAGTTGCTGTTATATTTTGGACTATTTCCGAGAAATTAAGAGG 350
QY 1350 AGAGCCATATCATATCTTCAAGCCCTTATCTTACCCCTGATATTAATCTGACAAAGA 1409
DB 351 GAGGCAATATCTGATCTTCAAGCCCTTATCTTACCCCTGATATTAATCTGACAAAGA 410
QY 1410 TGGTATATATATATGAAATTTCCATTAAACACTTGCATGATGATCTTCTTAC 1469
DB 411 TGGTATATATATATGAAATTTCCATTAAACACTTGCATGATGATCTTCTTAC 470
QY 1470 CAACCTTGTGTTTCTTATATCTGACCAAGTATCTATTTG-AAAGTGAACCTTGCTCCAA 1528
DB 471 CAACCTTGTGTTTCTTATATCTGACCAAGTATCTATTTG-AAAGTGAACCTTGCTCCAA 530
QY 1529 AATTGATGATATTTGATGCTGTTGCGAAGGACAGTGAAGAGAACTGACAGACCA 1588
DB 531 AATTGATGATATTTGATGCTGTTGCGAAGGACAGTGAAGAGAACTGATGATGACCA 590
QY 1589 TTCTAGTCAGAAATGAAATATCAATTAATGAATTTGACCTGTGAAACCTGCGCAGA 1648
DB 591 TTTTGTGCAAAAATGAAA--TATAATTAAGATGACCTTGACCTTGAAGCAGCAGACA 648
QY 1649 GCCTAACCTTCAGTTCAACTTTCA 1672
DB 649 ACATGACCTTCAGCTTAACCTTTCA 672

RESULT 13
DA361315 548 bp mRNA linear EST 05-NOV-2005
LOCUS DA361315 BRSTN2 Homo sapiens cDNA clone BRSTN2007643 5', mRNA
DEFINITION
sequence.
DA361315
DA361315.1 GI:80797290
VERSION
KEYWORDS
EST.
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
1 (bases 1 to 548)

AUTHORS
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Koshida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanabe, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
TITLE
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
JOURNAL
Genome Res. 16 (1), 55-65 (2006)
PUBMED
16344560
COMMENT
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
MEDO human CDNA project (New Energy and Industrial Technology
Developmental Organization, Japan): cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
FEATURES
source
1..548
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRSTN2007643"
/issue_type="subthalamic nucleus"
/clone_lib="BRSTN2"
/note="Vector: pME185FL3"

ORIGIN
Query Match 22.9%; Score 398.4; DB 9; Length 548;
Best Local Similarity 86.9%; Pred. No. 9.2e-108;
Matches 438; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
1 ATGCCCTTCCATGTGGAAGAGCTGTAGCTATTATCTCTTCACTCCTTAATTTCG 60
45 ATGGCTTTCATGTGGAAGAGCTGTAGCTATTATCTCTTCACTCCTTAATTTCG 104
61 GTTGAATATGGGCTGCATGGAACCAAAACAGGGGCAACCCAGAGGCGAGTGA 120
105 GTTGAATATGGGCTGCATGGAACCAAAACAGGGGCAACCCAGAGGCGAGTGA 164
121 GCCATCATATGTCGGGGCGGTGACATTTGTTGTTGTTGTTTACATGACAGCC 180
165 GCCATCATATGTTGTTGCGGAGATTTGTTTATTTGTTGTTGTTTACATGACAGCT 224
181 ACCTGGGTTGGAAGAGCTATCATTAATGGACAGCAAGACAGTATGGGCGAGTTGT 240
225 ACCTGGGTTGGAAGAGCTATCATTAATGGACAGCAAGACAGTATGGGCGAGTTAT 284
241 GGCTGCTTGGGCTCATGACCACTTGATTTCTGAGCTAATTTAGTGGTCTG 300
285 GGCTGCTTGGGCTCATGACCACTTGATTTCTGAGCTAATTTAGTGGGCTG 344
301 TTTTTCGAAACCTATGGCTTCAAGGATATGATATGTTAGACCATTCAAACAG 360
345 TTTTTCGAAACCTATGCTTCAAGGATATGATATGTTAGACCATTCAGCAA 404
361 ATCTATGGAAGAGCTATGGGCTGCTTTCATCCCTGACATGATGGAGAGATTC 420
405 ATCTATGGAAGAGCTATGGGCTGCTTTCATCCCTGACATGATGGAGAGATTC 464
421 TGGGCTGACAGCAATTTCTCTGCTATGAGGGCCACCATGACGTTGATGAT 480
465 TGGGCTGACAGCAATTTCTCTGCTTGGGAGCCACCATGACGTTGATGAT 524
481 GTGAACATATGGTCAATGCTCT 504

Db 525 ATGCACATTTCTGTCATCATCTCT 548
RESULT 14
DN991044
LOCUS
DN991044 576 bp mRNA linear EST 17-MAY-2005
DEFINITION
TC112868 Human adult whole brain, large insert, pcMV expression
library Homo sapiens cDNA clone TC112868 5' similar to Homo sapiens
soluble carrier family 5 (choline transporter), member 7 (SLC5A7),
mRNA sequence.
ACCESSION
DN991044
VERSION
DN991044.1 GI:66250875
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 576)
Birkett, C., Cho, Y., Gau, Y., Hamer, R., Kelly, S., Kovacs, K., Liu, L.,
Liu, X., Porter, J., Sachs, A., Shu, Y., Sun, Z., Wong, J., Wu, M.,
Zhang, X., Jay, G. and He, W.
High-throughput cloning of full-length human cDNAs directly from
cDNA libraries optimized for large and rare transcripts
Unpublished (2005)
Contact: Kovacs, KF
High Throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cDNA@origene.com
This EST submission is part of an on-going human full-length
cloning project at Origene Technologies, Inc.
Please contact Origene for access.
Origene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.origene.com
Seq primer: pcMV6 5prime forward vector primer, Origene
Technologies Inc.
FEATURES
source
1..576
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC112868"
/issue_type="Whole brain"
/clone_lib="Human adult whole brain, large insert, pcMV
expression library"
/note="Organ: Brain; Vector: pcMV6-XL5; Site:1: ECoRI;
Site:2: XhoI/SalI compatible end ligatio; Oligo-dT primed
reverse transcription optimized for large and GC rich mRNA
transcripts, cDNA size selection, optimized ligation for
large inserts into mammalian expression vector, random
clones selected for end sequence verification of
full-length genes"

ORIGIN
Query Match 22.5%; Score 391.8; DB 9; Length 576;
Best Local Similarity 87.4%; Pred. No. 9.2e-106;
Matches 429; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
1 ATGCCCTTCCATGTGGAAGAGCTGTAGCTATTATCTCTTCACTCCTTAATTTCG 60
85 ATGGCTTTCATGTGGAAGAGCTGTAGCTATTATCTCTTCACTCCTTAATTTCG 144
61 GTTGAATATGGGCTGCATGGAACCAAAACAGGGGCAACCCAGAGGCGAGTGA 120
145 GTTGAATATGGGCTGCATGGAACCAAAACAGGGGCAACCCAGAGGCGAGTGA 204
121 GCCATCATATGTCGGGGCGGTGACATTTGTTGTTGTTTACATGACAGCC 180


```
Db 181 |ACCTGGGTTGGAGGAGGCTACATCATGAGACAGAGAGAGAGTATGAGGCTAGTGT 240
Qy 241 |GATCAGCTGGGCTCATGACCACTAGTGAATATCTGAGCTAATTTTAGTGTCTG 300
Db 241 |GCTCAGCTGGGCTCATGACCACTAGTGAATATCTGAGCTAATTTTAGTGTCTG 300
Qy 301 |TTTTTGGGAAACCTATGCGTTCCAAAGGATATGACTATGTTAGAACCATTCACAG 360
Db 301 |TTTTTGGGAAACCTATGCGTTCCAAAGGATATGACTATGTTAGAACCATTCACAG 360
Qy 361 |ATCTATGAAAAGCGATGAGTGGGCTGCTCTTCAACCTCCGCACTGATGAGAGATGTTCT 420
Db 361 |ATCTATGAAAAGCGATGAGTGGGCTGCTCTTCAACCTCCGCACTGATGAGAGATGTTCT 420
Qy 421 |TGGGCTGAGCAAAATTTCTCTGCTCATTTAGGGGCGACCATCAGCGTATCATTTAGTGTGAT 480
Db 421 |TGGGCTGAGCAAAATTTCTCTGCTCATTTAGGGGCGACCATCAGCGTATCATTTAGTGTGAT 480
Qy 481 |GTGAACATATCGGCTATGCTCTGCACTCATTTGCCATTTCTTATACCTTAGTGGGTGGG 540
Db 481 |GTGAACATATCGGCTATGCTCTGCACTCATTTGCCATTTCTTATACCTTAGTGGGTGGG 540
Qy 541 |CTCTACTCTGAGCATATATCTGATGTTGTCAGCTATTTCTGCAATTTTATAGACTGTGG 600
Db 541 |CTCTACTCTGAGCATATATCTGATGTTGTCAGCTATTTCTGCAATTTTATAGACTGTGG 600
Qy 601 |ATCAGTGTCCCTTTTGCCCTGTACATCCTGAGTCAACGAGCATGGATTCACAGCTGTG 660
Db 601 |ATCAGTGTCCCTTTTGCCCTGTACATCCTGAGTCAACGAGCATGGATTCACAGCTGTG 660
Qy 661 |CATGCTAAATACCAAGTCTCCGCTGGGAGCAATGAAATCAGTTGAAGTTACACTTGG 720
Db 661 |CATGCTAAATACCAAGTCTCCGCTGGGAGCAATGAAATCAGTTGAAGTTACACTTGG 720
Qy 721 |CTTGATATTTTCTGTTATGATGCTGGGTGGAATCCCATGGCAAGCCCTACTTCCAGAG 780
Db 721 |CTTGATATTTTCTGTTATGATGCTGGGTGGAATCCCATGGCAAGCCCTACTTCCAGAG 780
Qy 781 |GTCTCTCTTCATCCTTCAGCACCCTATGCTCAGTATGCTCTTCTGAGCTTTTGGG 840
Db 781 |GTCTCTCTTCATCCTTCAGCACCCTATGCTCAGTATGCTCTTCTGAGCTTTTGGG 840
Qy 841 |TGCTCTGATGAGCTCTACCCGCCCATATGCAATGAGAGCTATTTGAGCTTCCACAGCTGG 900
Db 841 |TGCTCTGATGAGCTCTACCCGCCCATATGCAATGAGAGCTATTTGAGCTTCCACAGCTGG 900
Qy 901 |AACGAGACGCTAGGGGATCCAGATCCAGACCTAAGAGAGAGAGACATGATTC 960
Db 901 |AACGAGACGCTAGGGGATCCAGATCCAGACCTAAGAGAGAGAGACATGATTC 960
Qy 961 |CCGATCGTTCTGACAGTACTCTGCCCTGTGTACATCTCTTCTTGGGCTGTGTGT 1020
Db 961 |CCGATCGTTCTGACAGTACTCTGCCCTGTGTACATCTCTTCTTGGGCTGTGTGT 1020
Qy 1021 |TCAGCTGCTGATATGCTCTCACTGATCTGTCATCTGTCCGCAAGTCTATGTTTGT 1080
Db 1021 |TCAGCTGCTGATATGCTCTCACTGATCTGTCATCTGTCCGCAAGTCTATGTTTGT 1080
Qy 1081 |CGGAATATCTACAGCTTCTCTTCAGCAAAATGATCAGACAAAGAAATGTTGGGCT 1140
Db 1081 |CGGAATATCTACAGCTTCTCTTCAGCAAAATGATCAGACAAAGAAATGTTGGGCT 1140
Qy 1141 |ATGAGATCATCTGCTGTGTGTTGGAGCATCTGCAACGAGCATGGCTTGTCTGACGAG 1200
Db 1141 |ATGAGATCATCTGCTGTGTGTTGGAGCATCTGCAACGAGCATGGCTTGTCTGACGAG 1200
Qy 1201 |ACTGTGTATGGGCTCTGGATCCTGAGCTCTGCACTTGTCTCATCATCTTCCACAG 1260
Db 1201 |ACTGTGTATGGGCTCTGGATCCTGAGCTCTGCACTTGTCTCATCATCTTCCACAG 1260
Qy 1261 |CTGCTGTGTATCTCTTCATCAAGAGCAACACTTATGGGGAGTTGCTGTTATAT 1320
Db 1261 |CTGCTGTGTATCTCTTCATCAAGAGCAACACTTATGGGGAGTTGCTGTTATAT 1320
```

```
Db 1261 |CTGCTGTGTATCTCTTCATCAAGAGCAACACTTATGGGGAGTTGCTGTTATAT 1320
Qy 1321 |TTTGGACTATTTCTGAGAAATTAAGAGAGAGAGCATATCTATCTTACAGCCCTTAATC 1380
Db 1321 |TTTGGACTATTTCTGAGAAATTAAGAGAGAGAGCATATCTATCTTACAGCCCTTAATC 1380
Qy 1381 |TTTCACTGTGTTATTAATCTGACAGAGATGATATACAAATCAGAGTTCCCATTTAA 1440
Db 1381 |TTTCACTGTGTTATTAATCTGACAGAGATGATATACAAATCAGAGTTCCCATTTAA 1440
Qy 1441 |ACTCTCTCAATGTTAATCTCATTTTCAACAACTTGTGTTCTTATCTAGCCAAAGAT 1500
Db 1441 |ACTCTCTCAATGTTAATCTCATTTTCAACAACTTGTGTTCTTATCTAGCCAAAGAT 1500
Qy 1501 |CTATTTGAAATGGAACCTTCTCCAAATTAAGTATGATTTGATGCTGTGTGCAAG 1560
Db 1501 |CTATTTGAAATGGAACCTTCTCCAAATTAAGTATGATTTGATGCTGTGTGCAAG 1560
Qy 1561 |CACAGTGAAGAACATGAGACAGACATTTCTAGTCAAGAAATATCAAAATTAAT 1620
Db 1561 |CACAGTGAAGAACATGAGACAGACATTTCTAGTCAAGAAATATCAAAATTAAT 1620
Qy 1621 |GAACCTTGACCTGTGAAACCTGAGAGCCTAACCTCAGTTCAACTTCAACCAATAG 1680
Db 1621 |GAACCTTGACCTGTGAAACCTGAGAGCCTAACCTCAGTTCAACTTCAACCAATAG 1680
Qy 1681 |GAGGCCCTCTTGAATGTTATTCAGTCCGAGGGGCTGAGGACTGAGATTAATTTCA 1740
Db 1681 |GAGGCCCTCTTGAATGTTATTCAGTCCGAGGGGCTGAGGACTGAGATTAATTTCA 1740
Qy 1741 |TGA 1743
Db 1741 |TGA 1743

RESULT 2
US-09-911-077A-23
; Sequence 23, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:00808
; CURRENT APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-09-911-077A-23

Query Match 100.0%; Score 1743; DB 3; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 |ATGCCCTTTCATGTTGAGAGAGCTGTAGCTATTAATCTCTTCTTCACTCTTATATTTCTG 60
Db 1 |ATGCCCTTTCATGTTGAGAGAGCTGTAGCTATTAATCTCTTCTTCACTCTTATATTTCTG 60
Qy 61 |GTTGGAATATATGGGCTGCAATGGAATAACCAAAACGCGGCAACCCAGAGAGAGCGAGTAA 120
Db 61 |GTTGGAATATATGGGCTGCAATGGAATAACCAAAACGCGGCAACCCAGAGAGAGCGAGTAA 120
Qy 121 |GCCATCATATAGTCGGGGGCGGTGACATGTTGTTGTTGTTGTTTACCATGACAGCC 180
Db 121 |GCCATCATATAGTCGGGGGCGGTGACATGTTGTTGTTGTTGTTTACCATGACAGCC 180
```

```

QY 181 ACCTGGTGGAGAGGCTACATCAATGGGAGCAGAGAGTGTATGGCCAGGTTGT 240
DB 181 ACCTGGTGGAGAGGCTACATCAATGGGAGCAGAGAGTGTATGGCCAGGTTGT 240
QY 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATCTCTAGTCTAAATTTTAAAGTGTCTG 300
DB 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATCTCTAGTCTAAATTTTAAAGTGTCTG 300
QY 301 TTTTTCGCAAACTATGCGTTCAGAGGATATGTACTATGTTTAAACCATTTCAACAG 360
DB 301 TTTTTCGCAAACTATGCGTTCAGAGGATATGTACTATGTTTAAACCATTTCAACAG 360
QY 361 ATCTATGAAAGGAGCTGGGCTGCTCTTCATCCCTGCACTGATGAGAGAGATGTTTC 420
DB 361 ATCTATGAAAGGAGCTGGGCTGCTCTTCATCCCTGCACTGATGAGAGAGATGTTTC 420
QY 421 TGGGCTGAGCAATTTCTCTGCAATTAAGGAGCCACCATGAGCGTGTATCATTTGATGAT 480
DB 421 TGGGCTGAGCAATTTCTCTGCAATTAAGGAGCCACCATGAGCGTGTATCATTTGATGAT 480
QY 481 GTGAAATATCGGTCTATGCTCTGCACTATGCAATTCCTTATACCTTATGAGGTTGGG 540
DB 481 GTGAAATATCGGTCTATGCTCTGCACTATGCAATTCCTTATACCTTATGAGGTTGGG 540
QY 541 CTCTACTCTGAGCATATATGATGATGTTGTCAGACTATTCGCAATTTTATAGAGCTGTGG 600
DB 541 CTCTACTCTGAGCATATATGATGATGTTGTCAGACTATTCGCAATTTTATAGAGCTGTGG 600
QY 601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGCACTGAGCAGCAGATCGGATTCACAGCTGTG 660
DB 601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGCACTGAGCAGCAGATCGGATTCACAGCTGTG 660
QY 661 CATGCTAATATACAGAGTCCCTGCTGAGGAGAACATTTGATCACTGATGATGATGATGAT 720
DB 661 CATGCTAATATACAGAGTCCCTGCTGAGGAGAACATTTGATCACTGATGATGATGATGAT 720
QY 721 CTGTGATATTTTCTGTATTTGATGCTGGGAGAAATCCATGGAGCCTTACTTCCAGAGG 780
DB 721 CTGTGATATTTTCTGTATTTGATGCTGGGAGAAATCCATGGAGCCTTACTTCCAGAGG 780
QY 781 GTCTCTCTTCACTCTCAGCAGCACTATGCTCAGAGTATGCTCTTCTGAGCTTTTGGG 840
DB 781 GTCTCTCTTCACTCTCAGCAGCACTATGCTCAGAGTATGCTCTTCTGAGCTTTTGGG 840
QY 841 TGCCTGTGATGCTCTACCCCGCATATGATGAGCTATTTGAGCTTCCACAGACTGG 900
DB 841 TGCCTGTGATGCTCTACCCCGCATATGATGAGCTATTTGAGCTTCCACAGACTGG 900
QY 901 AACCAAGCTGCTAGGGGATATCCAGATCCCAAGATTAAGGAGAGCAGATGATTTCTC 960
DB 901 AACCAAGCTGCTAGGGGATATCCAGATCCCAAGATTAAGGAGAGCAGATGATTTCTC 960
QY 961 CCGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 CCGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 TGAAGCTGCTGCTATGCTCTGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 TGAAGCTGCTGCTATGCTCTGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 CCGAATATATCAGAGCTTTCTCAGACAAATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 CCGAATATATCAGAGCTTTCTCAGACAAATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 ATGAGATATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1141 ATGAGATATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260

```

```

QY 1261 CTGCTCTGTGTACTCTTATCAAGAACCAACTTATGGGCACTTCTGCTTATAT 1320
DB 1261 CTGCTCTGTGTACTCTTATCAAGAACCAACTTATGGGCACTTCTGCTTATAT 1320
QY 1321 TTTGAGCTATTCCTGAGAAATTAAGGAGAGCCATATCTATCTTGCAGCCCTTAATC 1380
DB 1321 TTTGAGCTATTCCTGAGAAATTAAGGAGAGCCATATCTATCTTGCAGCCCTTAATC 1380
QY 1381 TTCTACCTGCTTATTAATCTGACAAAGATGATATATCAATCAGAGGTTCCATTAA 1440
DB 1381 TTCTACCTGCTTATTAATCTGACAAAGATGATATATCAATCAGAGGTTCCATTAA 1440
QY 1441 ACTCTCTCCATGATTAATCTTATCAATTCCTTATCAATTCCTTATCAATTCCTTAT 1500
DB 1441 ACTCTCTCCATGATTAATCTTATCAATTCCTTATCAATTCCTTATCAATTCCTTAT 1500
QY 1501 CTATTTGAAAGTGAAACCTTGCCTCCAAATTAATGATGATGATGATGATGATGATGAT 1560
DB 1501 CTATTTGAAAGTGAAACCTTGCCTCCAAATTAATGATGATGATGATGATGATGATGAT 1560
QY 1561 CACAGTGAAGAGAACATGAGCAAGACATTCATGTCAGAAATGAAATATCAATTAAT 1620
DB 1561 CACAGTGAAGAGAACATGAGCAAGACATTCATGTCAGAAATGAAATATCAATTAAT 1620
QY 1621 GAACCTTGCACTGTGAAACCTGCGCAGAGCCTTAACTTCAATTCCTTCAATTAAG 1680
DB 1621 GAACCTTGCACTGTGAAACCTGCGCAGAGCCTTAACTTCAATTCCTTCAATTAAG 1680
QY 1681 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 TGA 1743
DB 1741 TGA 1743

```

```

RESULT 3
US-10-724-806-3
; Sequence 3, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:00805
; CURRENT APPLICATION NUMBER: US/10/724,806
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-10-724-806-3

```

```

Query Match 100.0%; Score 1743; DB 9; Length 1743;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTTTTCATGAGAGAGCTAGTATATCTCTTCACTCTTATATTTCTG 60
DB 1 ATGCTTTTCATGAGAGAGCTAGTATATCTCTTCACTCTTATATTTCTG 60
QY 61 GTTGAATATGGGCTGCTGAGAAACCAAAACAGGGCAACCCAGAGGCGCAGTGA 120
DB 61 GTTGAATATGGGCTGCTGAGAAACCAAAACAGGGCAACCCAGAGGCGCAGTGA 120

```

121 GCCATCATAGTCGGGGCCGTGACATGTTGTTGTTGTTGTTGTTTACCATGACAGC 180
121 GCCATCATAGTCGGGGCCGTGACATGTTGTTGTTGTTGTTTACCATGACAGC 180
181 ACCTGGGTTGGAGAGGCTACATCAATGGGACAGAGAGCAGTGTATGGCCAGGTTGT 240
181 ACCTGGGTTGGAGAGGCTACATCAATGGGACAGAGAGCAGTGTATGGCCAGGTTGT 240
241 GGTACGCTTGGGCTCATGCAACCATGGAATTTCTTGAGCTAATTTTAAAGTGTCTG 300
241 GGTACGCTTGGGCTCATGCAACCATGGAATTTCTTGAGCTAATTTTAAAGTGTCTG 300
301 TTTTTCGGAAAACCTATGGGTTTCCAAAGGATATGTGACTATGTAGACCATTTCAAACG 360
301 TTTTTCGGAAAACCTATGGGTTTCCAAAGGATATGTGACTATGTAGACCATTTCAAACG 360
361 ATCTATGGAAGGAGATGGGTGGGCTGCTTCCATCCCTGCACTGATGGGAGAGATGTTT 420
361 ATCTATGGAAGGAGATGGGTGGGCTGCTTCCATCCCTGCACTGATGGGAGAGATGTTT 420
421 TGGGCTGAGCAATTTTCTGCAATTAAGGGCCACATCAGGCTGATCATGTATGTGAT 480
421 TGGGCTGAGCAATTTTCTGCAATTAAGGGCCACATCAGGCTGATCATGTATGTGAT 480
481 GTGAACATATCGGTCATTTGCTGCACTCATTTGCCATTCTTTAATCCCTAGTGGGTGG 540
481 GTGAACATATCGGTCATTTGCTGCACTCATTTGCCATTCTTTAATCCCTAGTGGGTGG 540
541 CTCTATCTGTTGGGCTATATCTGATGTGTCAGCTATTTCTGATTTTATAGACTGTGG 600
541 CTCTATCTGTTGGGCTATATCTGATGTGTCAGCTATTTCTGATTTTATAGACTGTGG 600
601 ATCACTGTCCTTTTGGCCCTGTACATCTGTGAGTCAACGATCGGATTTCAACGCTGG 660
601 ATCACTGTCCTTTTGGCCCTGTACATCTGTGAGTCAACGATCGGATTTCAACGCTGG 660
661 CATGCTAAATACCAAGATCCCTGGCTGGGACCATTTGAATCAAGTGTACACCTGG 720
661 CATGCTAAATACCAAGATCCCTGGCTGGGACCATTTGAATCAAGTGTACACCTGG 720
721 CTGATATATTTCTGTTATGTAGTCTGGGTGGAAATCCCATGGCAAGCCATCTTCCAGAG 780
721 CTGATATATTTCTGTTATGTAGTCTGGGTGGAAATCCCATGGCAAGCCATCTTCCAGAG 780
781 GTCTCTCTTCACTCTCAAGCACTATGCTCAGGTAAGTCTCTTCTGCACTTTTGGG 840
781 GTCTCTCTTCACTCTCAAGCACTATGCTCAGGTAAGTCTCTTCTGCACTTTTGGG 840
841 GTCTCTCTTCACTCTCAAGCACTATGCTCAGGTAAGTCTCTTCTGCACTTTTGGG 900
841 GTCTCTCTTCACTCTCAAGCACTATGCTCAGGTAAGTCTCTTCTGCACTTTTGGG 900
901 AACGAGACTGCTTACGGGATCCAGATCCCAAGATCCCAAGAGAGAGAGATGATTC 960
901 AACGAGACTGCTTACGGGATCCAGATCCCAAGATCCCAAGAGAGAGAGATGATTC 960
961 CCGATGCTTCTGACAGTACCTTGCCTGTGTATCATCTCTTCTTGGGCTGTGGTGT 1020
961 CCGATGCTTCTGACAGTACCTTGCCTGTGTATCATCTCTTCTTGGGCTGTGGTGT 1020
1021 TCAGCTGCTGTATGCTCTCAGCTGACTGCTGCTCATCTGTGCGAGATTTATGTTTGT 1080
1021 TCAGCTGCTGTATGCTCTCAGCTGACTGCTGCTCATCTGTGCGAGATTTATGTTTGT 1080
1081 CGGAATATCTACAGCTTCTCTGAGCAAAATGCATCGAGAGAGAAATGTGTGGGTC 1140
1081 CGGAATATCTACAGCTTCTCTGAGCAAAATGCATCGAGAGAGAAATGTGTGGGTC 1140
1141 ATGAGATCATCTGCTTGTGTGCGAGCATCTGCAACGAGCATGCTTTGTGTCAGAG 1200
1141 ATGAGATCATCTGCTTGTGTGCGAGCATCTGCAACGAGCATGCTTTGTGTCAGAG 1200

1201 ACTGTATAGGAGCTCTGATCTGAGCTCTGACCTTGTCTACATCATCATCTTCCACAG 1260
1201 ACTGTATAGGAGCTCTGATCTGAGCTCTGACCTTGTCTACATCATCATCTTCCACAG 1260
1261 CTGCTCTGTATCTTTTATCATTAAGAGAACCAACTTATGGGGAGCTTGTGTATATT 1320
1261 CTGCTCTGTATCTTTTATCATTAAGAGAACCAACTTATGGGGAGCTTGTGTATATT 1320
1321 TTTGACTATTTCTGAGAAATTAAGAGAGAGAGCCATATCTATCTTACAGCCCTTAATC 1380
1321 TTTGACTATTTCTGAGAAATTAAGAGAGAGAGCCATATCTATCTTACAGCCCTTAATC 1380
1381 TTTGACTATTTCTGAGAAATTAAGAGAGAGAGCCATATCTATCTTACAGCCCTTAATC 1440
1381 TTTGACTATTTCTGAGAAATTAAGAGAGAGAGCCATATCTATCTTACAGCCCTTAATC 1440
1441 ACTCTCTGAGTATCTGATCTTCTTCAAACTTTGTTCTTATCTTACAGCCAGAT 1500
1441 ACTCTCTGAGTATCTGATCTTCTTCAAACTTTGTTCTTATCTTACAGCCAGAT 1500
1501 CTATTTGAAAGTGAACCTTCCCTCCAAATTAAGATGATTTGATGCTGTTGCGAAG 1560
1501 CTATTTGAAAGTGAACCTTCCCTCCAAATTAAGATGATTTGATGCTGTTGCGAAG 1560
1561 CACAGTGAAGAGACATGAGACAGACCATTTCTAGTCAAGAAATGAATAATTAAT 1620
1561 CACAGTGAAGAGACATGAGACAGACCATTTCTAGTCAAGAAATGAATAATTAAT 1620
1621 GAACCTTGACCTGGAACCTTCCGAGAGCCATACCCCTGATCACTTACCAATTAAG 1680
1621 GAACCTTGACCTGGAACCTTCCGAGAGCCATACCCCTGATCACTTACCAATTAAG 1680
1681 GAGGCCCTCTTGAATGATTTCCAGTCCGAGAGGGCTGGGACTGAAGATTAATCA 1740
1681 GAGGCCCTCTTGAATGATTTCCAGTCCGAGAGGGCTGGGACTGAAGATTAATCA 1740
1741 TGA 1743
1741 TGA 1743

RESULT 4
US-10-724-806-23
; Sequence 23, Application US-10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:00805
; CURRENT APPLICATION NUMBER: US/10/724,806
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-10-724-806-23

Query Match 100.0%; Score 1743; DB 9; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCTTTTCCATGTTGAAGAGACTGATGATCTTATCTTCTTATATTTTCTG 60
1 ATGCTTTTCCATGTTGAAGAGACTGATGATCTTATCTTCTTATATTTCTG 60

```

OY 61 GTTGAATATGAGCTCATGAAAAACAAAACAGCGCAACCCAGAGGCGCAGTGA 120
DB 61 GTTGAATATGAGCTCATGAAAAACAAAACAGCGCAACCCAGAGGCGCAGTGA 120
OY 121 GCCATCATAGTGGGGGCGGTGACATTTGTTGGTTGGTTTAAACATGACAGCC 180
DB 121 GCCATCATAGTGGGGGCGGTGACATTTGTTGGTTGGTTTAAACATGACAGCC 180
OY 181 ACCGTGGTTGAGAGAGCTACATCAATGAGACAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 ACCGTGGTTGAGAGAGCTACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
OY 241 GGTCTAGCTTGGAGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTGTCTG 300
DB 241 GGTCTAGCTTGGAGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTGTCTG 300
OY 301 TTTTTCGAAAACCTATGCGTTCCAAAGGATATGTGACTATGTAGACCATCAACAG 360
DB 301 TTTTTCGAAAACCTATGCGTTCCAAAGGATATGTGACTATGTAGACCATCAACAG 360
OY 361 ATCTATGAGAAAGGAGTGGGTGGGCTCTTCAATCCCTGCACTGATGAGAGAGAGTTC 420
DB 361 ATCTATGAGAAAGGAGTGGGTGGGCTCTTCAATCCCTGCACTGATGAGAGAGAGTTC 420
OY 421 TGGGCTGAGCAATTTTCTGCTGATAGAGGCGCACCATGATGATGATGATGAT 480
DB 421 TGGGCTGAGCAATTTTCTGCTGATAGAGGCGCACCATGATGATGATGATGATGAT 480
OY 481 GTGAACATATCGGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 GTGAACATATCGGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
OY 541 CTCTACTCTGTGGCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 CTCTACTCTGTGGCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 600
OY 601 ATCAGAGTCCCTTTTCCCTGTCACATCTGACAGTACCGACATCGAGATTCACAGCTGTG 660
DB 601 ATCAGAGTCCCTTTTCCCTGTCACATCTGACAGTACCGACATCGAGATTCACAGCTGTG 660
OY 661 CATGCTAAATACCAAGTCCCTGCTGAGGAAACCATTTGAATCAGTTGAGTCAACCTGTG 720
DB 661 CATGCTAAATACCAAGTCCCTGCTGAGGAAACCATTTGAATCAGTTGAGTCAACCTGTG 720
OY 721 CTGTAATATTTTCTGTTATGATGCTGGGTGAAATCCCATGAGACCTTCCAGAGG 780
DB 721 CTGTAATATTTTCTGTTATGATGCTGGGTGAAATCCCATGAGACCTTCCAGAGG 780
OY 781 GTCTCTCTTCAATCCCTGACGACCTATGCTGAGGTGCTGCTGCTGCTGCTGCTGCTG 840
DB 781 GTCTCTCTTCAATCCCTGACGACCTATGCTGAGGTGCTGCTGCTGCTGCTGCTGCTG 840
OY 841 TGCCTGTGATGCTTACCCGACATATGATGAGAGTATTTGAGCTTCCACAGCTGG 900
DB 841 TGCCTGTGATGCTTACCCGACATATGATGAGAGTATTTGAGCTTCCACAGCTGG 900
OY 901 AACCAACCTGCTTACCGGATATCCAGATCCCAAGACTTAAGAGAGAGAGAGAGATTC 960
DB 901 AACCAACCTGCTTACCGGATATCCAGATCCCAAGACTTAAGAGAGAGAGAGAGATTC 960
OY 961 CGGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 CGGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
OY 1021 TCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1021 TCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
OY 1081 CGGAATATCTACAGGCTTTCCTTCAAGCAAAATGATCAGACAAAGAAATGTGTGGGTC 1140
DB 1081 CGGAATATCTACAGGCTTTCCTTCAAGCAAAATGATCAGACAAAGAAATGTGTGGGTC 1140

```

```

OY 1141 ATGAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1141 ATGAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
OY 1201 ACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
DB 1201 ACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
OY 1261 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1261 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
OY 1321 TTTGACATATTCCTGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1321 TTTGACATATTCCTGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
OY 1381 TTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1381 TTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
OY 1441 ACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1441 ACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
OY 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTAAGATGATGATGATGATGATGATGATG 1560
DB 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTAAGATGATGATGATGATGATGATGATG 1560
OY 1561 CACAGTGAAGAGAACTGAGAACATCTTCTGATGATGATGATGATGATGATGATGATG 1620
DB 1561 CACAGTGAAGAGAACTGAGAACATCTTCTGATGATGATGATGATGATGATGATGATG 1620
OY 1621 GAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 1621 GAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
OY 1681 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
DB 1681 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
OY 1741 TGA 1743
DB 1741 TGA 1743

RESULT 5
US-09-911-077A-5
; Sequence 5, Application US/0911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELEY, RANDY D.
; APPLICANT: APPARUNDABAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4904
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; NAME/KEY: CDS
; LOCATION: (224)..(1966)
US-09-911-077A-5

Query Match 92.7%; Score 1616.6; DB 3; Length 4904;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

OY 1 ATGCTTTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60

```

```
Db 224 ATGCTTTCATGTAAGAGACTAGTAGGATTATCTGTTCTACCTCTTATTTTCG 283
Qy 61 GTTGAATATGGGCTGCATGAGAAAACAAAACAGCGGCACCCAGAGAGCGAGTGA 120
Db 284 GTTGAATATGGGCTGCATGAGAAAACAAAACAGCGGTATAGAGAAAGCAGCGGA 343
Qy 121 GCCATCATAGTCGGGGGCGGTGACATTTGTTGTTGTTGTTTACATGACGCC 180
Db 344 GCCATCATAGTTGGGGCGGAGACATTTGTTGTTGTTGTTTACATGACGCC 403
Qy 181 ACCTGGTTGGAGAGGCTACATCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 404 ACCTGGTTGGAGAGGCTACATCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 463
Qy 241 GGTCTAGCTTGGGCTCATGACCCATTTGATATTTCTGAGTCTAATTTTGGTGTG 300
Db 464 GGTCTAGCTTGGGCTCATGACCCATTTGATATTTCTGAGTCTAATTTTGGTGTG 523
Qy 301 TTTTGGGAAACCTATGAGGTTCCAAAGGATATGATCATATGTTAGACCATTTCAAACG 360
Db 524 TTTTGGGAAACCTATGAGGTTCCAAAGGATATGATCATATGTTAGACCCGTTCAAACG 583
Qy 361 ATCTATGAGAAAGCGATGGGCTGCTCTCATCCCTGCACTGATGGAGAGATGTC 420
Db 584 ATCTATGAGAAAGCGATGGGCTGCTGCTCTCATCCCTGCACTGATGGAGAGATGTC 643
Qy 421 TGGGCTGAGCAATTTTCTGCAATTAAGGGGCGACCATGAGGATCATGATGATGAT 480
Db 644 TGGGCTGAGCAATTTTCTGCAATTAAGGGGCGACCATGAGGATCATGATGATGAT 703
Qy 481 GTGAACATATGGGTATGTTCTGTCATCTGATGTCATTTCTTATACCTAGTGGTGG 540
Db 704 GTGAACATATGGGTATGTTCTGTCATCTGATGTCATTTCTTATACCTAGTGGAGGG 763
Qy 541 CTCTACTCTGTCGCAATATCTGATGTTGTCAGACTATTCGATTTTATAGACGTGG 600
Db 764 CTCTACTCTGTCGCAATATCTGATGTTGTCAGACTATTCGATTTTATAGACGTGG 823
Qy 601 ATCAGTGTCCCTTTTGGCCCTGTACATCCTGTGACAGTCCAGCATCGGATTCACAGCTGTG 660
Db 824 ATCAGTGTCCCTTTTGGCCCTGTACATCCTGTGACAGTCCAGCATCGGATTCACAGCTGTG 883
Qy 661 CATGCTAATATACCAAGTCCCTGGGCTGGGAAACATTTGAATCAGTTGAAGTCTACACTGG 720
Db 884 CATGCTAATATACCAAGTCCCTGGGCTGGGAAACATTTGAATCAGTTGAAGTCTACACTGG 943
Qy 721 CTTGATATATTTCTGTTATTTAGTGGTGGATCCCATGAGCAAGCTACTTCCAGAG 780
Db 944 CTTGATATATTTCTGTTATTTAGTGGTGGATCCCATGAGCAAGCTACTTCCAGAG 1003
Qy 781 GTCTCTCTTCATCTCAGCAACCTATAGTCTGAGGATCTGCTCTTCTGAGCTTTTGGG 840
Db 1004 GTCTCTCTTCATCTCAGCAACCTATAGTCTGAGGATCTGCTCTTCTGAGCACTTTTGGG 1063
Qy 841 TGCCGTGATAGGCTTACCCGCAATAGCATAGAGCATATGAGCTTCCACACATGG 900
Db 1064 TGCCGTGATAGGCTTACCCGCAATAGCATAGAGCATATGAGCTTCCACACATGG 1123
Qy 901 AACCAAGACTGCTAGCGGATCCAGATCCCAAGACTAAGAGAGAGAGAGAGATGATTC 960
Db 1124 AACCAAGACTGCTAGCGGATCCAGATCCCAAGACTAAGAGAGAGAGAGAGATGATTC 1183
Qy 961 CCGATGCTTCTGAGTACTGCTGCCCTGTGATCATCTCTTCTTGGGCTGTGCTGT 1020
Db 1184 CCGATGCTTCTGAGTACTGCTGCCCTGTGATCATCTCTTCTTGGGCTGTGCTGT 1243
Qy 1021 TCAGTGTGCTGATGCTCCTCAGCTGATCTGCTGATCTGCTGCTGCTGCTGCTGCT 1080
Db 1244 TCAGTGTGCTGATGCTCCTCAGCTGATCTGCTGATCTGCTGCTGCTGCTGCTGCT 1303
Qy 1081 CGGAATATCTACAGCTTCTCTTCAAGCAAAATGATCAGAGAGAAATGTTGTGGGTC 1140
```

```
Db 1304 CGGAATATCTACAGCTTCTCTTCAAGCAAAATGATCAGAGAGAAATGTGTGGTC 1363
Qy 1141 ATGAGGATCACTGTGCTTGTGTGGAGCATCTGCAAGCCATGCTTTGTGACGAG 1200
Db 1364 ATGAGGATCACTGTGCTTGTGTGGAGCATCTGCAAGCCATGCTTTGTGACGAG 1423
Qy 1201 ACTGTGATAGGCTCTGTGTACCTGATCTGAGCTGACCTGTGCTATCATATCTTCCACAG 1260
Db 1424 ACTGTGATAGGCTCTGTGTACCTGATCTGAGCTGACCTGTGCTATCATATCTTCCACAG 1483
Qy 1261 CTGCTCTGTACTCTTCATCAAGAGAACCAACTTATGGGGCAGTTGCTGTTATAT 1320
Db 1484 CTGCTCTGTACTCTTCATCAAGAGAACCAACTTATGGGGCAGTTGCTGTTATAT 1543
Qy 1321 TTTGACTATTTCTGAGAAATTAATGAGAGAGAGCCATATCTATCTGACGCTTAATC 1380
Db 1544 TTTGACTATTTCTGAGAAATTAATGAGAGAGAGCCATATCTATCTGACGCTTAATC 1603
Qy 1381 TTTCAACCTGGTTATTAATCTGACAGAAATGTTATTAATCAATCAAGGTTCCATTTAA 1440
Db 1604 TTTCAACCTGGTTATTAATCTGACAGAAATGTTATTAATCAATCAAGGTTCCATTTAA 1663
Qy 1441 ACTCTCTCCAGTGTACTCTCATTTCTTACCAACATTTGTGTTCTTATCTAGCCAGTAT 1500
Db 1664 ACTCTCTCCAGTGTACTCTCATTTCTTACCAACATTTGTGTTCTTATCTAGCCAGTAT 1723
Qy 1501 CTATTTGAAAGTGAACTTGTGCTTCAAAATTTAGATGATTTGATGCTGTGTGCGAAG 1560
Db 1724 CTATTTGAAAGTGAACTTGTGCTTCAAAATTTAGATGATTTGATGCTGTGTGCGAAG 1783
Qy 1561 CACAGTGAAGAAATATGAGAACCACTCTGACAGAAATGAAATGAAATTAAT 1620
Db 1784 CACAGTGAAGAAATATGAGAACCACTCTGACAGAAATGAAATGAAATTAAT 1843
Qy 1621 GAACCTGACCTGTGAAACCTGCGAGAGCCTAACCTCAGTTCAACTTCAACCAATAG 1680
Db 1844 GAACCTGACCTGTGAAACCTGCGAGAGCCTAACCTCAGTTCAACTTCAACCAATAG 1903
Qy 1681 GAGGCTCTCTTATGATTTGATTTCCAGTCCGAGAGGCTGTGGAGCTGAGATTAATTA 1740
Db 1904 GAGGCTCTCTTATGATTTGATTTCCAGTCCGAGAGGCTGTGGAGCTGAGATTAATTA 1963
Qy 1741 TGA 1743
Db 1964 TGA 1966

RESULT 6
US-10-724-806-5
; Sequence 5, Application US/10724806
; Publication No. US20040248638A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:00805
; CURRENT APPLICATION NUMBER: US/10/724,806
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4904
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (224)..(1966)
US-10-724-806-5
Query Match 92.7%; Score 1616.6; DB 9; Length 4904;
```


NUMBER OF SEQ ID NOS: 27
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 9
 LENGTH: 1813
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (19) .. (1761)
 US-10-724-806-9

Query Match 78.9%; Score 1375; DB 9; Length 1813;
 Best Local Similarity 86.8%; Pred. No. 0;
 Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

```

QY 1 ATGCCCTTCCATGTGAAAGGACTGTAATATCTCTTCTACCTCCTTATATTTCTG 60
DB 19 ATGGCTTTCCATGTGAAAGGACTGTAATATCTCTTCTACCTCCTTATATTTCTG 78
QY 61 GTTGAATATGGGCTCATGAAAAACAAAACAGGCGCAACCCAGAGAGCGAGTGA 120
DB 79 GTTGAATATGGGCTCATGAAAAACAAAACAGGCGCAACCCAGAGAGCGAGTGA 138
QY 121 GCCATCATAGTCGGGGCCGTGAATTGTTGTTGTTGTTGTTTACATGACAGCC 180
DB 139 GCCATCATAGTTGGTGGCCGAGATATTTGTTTATGTTGGTTGATTTACATGACAGCT 198
QY 181 ACCGTGGTGGAGAGGCTACATCAATGGGACGAGAAAGCACTGATGAGGCGCAAGTTGT 240
DB 199 ACCGTGGTGGAGAGGCTACATCAATGGGACGAGAAAGCACTGATGATGACAGGTTAT 258
QY 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATTTCTGAGTCTAATTTTAAAGTGGCTG 300
DB 259 GGTCTAGCTTGGGCTCATGCAACCATTTGATTTCTGAGTCTAATTTTAAAGTGGCTG 318
QY 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTAATGTTAGACCCATTCAAACAG 360
DB 319 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTAATGTTAGACCCATTCAGCAA 378
QY 361 ATCTATGAAAGGCGATGGGTGGGCTGCTTCTTATCCCTGCACTGATGGAGAGATGTTT 420
DB 379 ATCTATGAAAGGCGATGGGTGGGCTGCTTCTTATCCCTGCACTGATGGAGAGATGTTT 438
QY 421 TGGGCTGAGCAATTTCTCTGATTAAGGGGCCACCATGAGGCTGATCAATGATGAT 480
DB 439 TGGGCTGAGCAATTTCTCTGATTAAGGGGCCACCATGAGGCTGATCAATGATGAT 498
QY 481 GTGAACATATCGGTCTATGCTCTGCACTATTCATTTCTTATACCTTATGAGGTGG 540
DB 499 ATGCAACATTTCTGATCACTCTGCACTATTCATTTGCACTGATGAGGTGGAGGG 558
QY 541 CTCTACTCTGTGGCATATACTGATGTTGTCAGCTATTTCTGCAATTTTATAGAGCTGTG 600
DB 559 CTCTATTTCTGTGGCTTACACTGATGTCGTTCACTCTTTTGAAGGCTGTG 618
QY 601 ATCAGAGTCCCTTTTGGCCGTCATCTGCACTCTGCACTGAGGAGTCCAGGATTCACAGCTG 660
DB 619 ATCAGAGTCCCTTTTGGCCGTCATCTGCACTCTGCACTGAGGAGTCCAGGATTCACAGCTG 678
QY 661 CATGCTAAATACAGAGTCCCTGCTGGAGAACCATTTGAATCAGTTGAAGTCTACACTG 720
DB 679 CATGCTAAATACAGAGTCCCTGCTGGAGAACCATTTGAATCAGTTGAAGTCTACACTG 738
QY 721 CTGTAATATTTCTGTTATGATGCTGGGTGAATCCCATGGAACCTTATCTTCCAGAG 780
DB 739 CTGTAATATTTCTGTTATGATGCTGGGTGAATCCCATGGAACCTTATCTTCCAGAG 798
QY 781 GTTCTCTCTTCACTCAGCACTATGCTCAGAGTCTGTCTTCTTGGAGCTTTTGG 840
DB 799 GTTCTCTCTTCTTCCAGCACTATGCTCAGAGTCTGTCTTCTTGGAGCTTTTGG 858
QY 841 TGCCTGTAAGGCTTACCCGCAATGATGAGCTATTTGAGCTTCCACAGACTGG 900

```

```

DB 859 TGCCTGTAAGGCTTACCCAGCACTATGAGGCTTGGAGCTGAGACTCAACAGACTGG 918
QY 901 AACCGAGCTGCTTACGGGATCCAGATCCCAAGACTTAAGAGGAGACAGATGATCTC 960
DB 919 AACCGAGCTGCTTACGGGATCCCAAGACTTAAGAGGAGACAGATGATGATCTC 978
QY 961 CCGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 979 CCAATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038
QY 1021 TCAAGCTGCTGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1039 TCTGCTGCTGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1098
QY 1081 CCGAATATCAGCTTCTTCAAGCAAAATGCTGAGCAAGGAAATGCTGAGGCTC 1140
DB 1099 CCGAATATCAGCTTCTTCAAGCAAAATGCTGAGCAAGGAAATGCTGAGGCTC 1158
QY 1141 ATGAGATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1159 ATGAGATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218
QY 1201 ACTGTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
DB 1219 ACTGTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1278
QY 1261 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 1279 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1338
QY 1321 TTTGCACTATTCCTGAGATTAATCTGAGAGAGAGCCATCTATCTGAGAGCCCTTATC 1380
DB 1339 TCTGCTGCTTCTCTGAGATTAATCTGAGAGAGAGCCATCTATCTGAGAGCCCTTATC 1398
QY 1381 TTTTACCTGCTTATTAATCTGAGCAAAATGCTATTAATCAATCAAGGTTCCATTTAA 1440
DB 1399 TTTTACCTGCTTATTAATCTGAGCAAAATGCTATTAATCAATCAAGGTTCCATTTAA 1458
QY 1441 ACTCTCTCATAGTATACCTTCTTATCAACATTTGCTGCTTCTTATCTAGCAAGTAT 1500
DB 1459 ACACTTCTCATAGTATACCTTCTTATCAACATTTGCTGCTTCTTATCTAGCAAGTAT 1518
QY 1501 CTATTTGAAGTGAACCTTGGCTCCAAATTAAGATGATTTGATCTGCTGCTGCTGCTG 1560
DB 1519 CTATTTGAAGTGAACCTTGGCTCCAAATTAAGATGATTTGATCTGCTGCTGCTGCTG 1578
QY 1561 CACAGTGAAGAGATGAGCAAGACATTTCTAGTCAAAATGAATCAAAATTAAT 1620
DB 1579 CACAGTGAAGAGATGAGCAAGACATTTCTAGTCAAAATGAATCAAAATTAAT 1638
QY 1621 GAATTTGCACTGCTGAGAACTTGGGCAAGCTTCAACCTTCACTTCACTTCACTTCA 1680
DB 1639 GAATTTGCACTGCTGAGAACTTGGGCAAGCTTCAACCTTCACTTCACTTCACTTCA 1698
QY 1681 GAGGCTTCTTGAATGATTTCAAGTCCGAGAGGCTTGGGCTGAGCTGAATTAATTA 1740
DB 1699 GAGGCTTCTTGAATGATTTCAAGTCCGAGAGGCTTGGGCTGAGCTGAATTAATTA 1758
QY 1741 TGA 1743
DB 1759 TGA 1761

```

RESULT 11
 US-11-005-030-1

; Sequence 1, Application US/11005030
 ; Publication No. US20050255495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aerssens, Jeroen
 ; APPLICANT: Athanasios, Maria
 ; APPLICANT: Brain, Carlos
 ; APPLICANT: Cohen, Nadine
 ; APPLICANT: Dain, Bradley

```

; APPLICANT: Denton, R. Rex
; APPLICANT: Judson, Richard S.
; APPLICANT: Ozdemir, Vural
; APPLICANT: Reed, Carol R.
; TITLE OF INVENTION: SLC5A7 Genetic Markers Associated with Age of Onset of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2300.0090001
; CURRENT APPLICATION NUMBER: US/11/005,030
; CURRENT FILING DATE: 2004-12-07
; PRIOR APPLICATION NUMBER: US 60/529,999
; PRIOR FILING DATE: 2003-12-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patent version 3.3
; SEQ ID NO 1
; LENGTH: 26323
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1973)..(1973)
; OTHER INFORMATION: n is 'g' or 'a'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2010)..(2010)
; OTHER INFORMATION: n is 't' or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2273)..(2273)
; OTHER INFORMATION: n is 'a' or 'g'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2417)..(2417)
; OTHER INFORMATION: n is 't' or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7671)..(7671)
; OTHER INFORMATION: n is 'g' or 't'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16520)..(16520)
; OTHER INFORMATION: n is 'c' or 't'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16546)..(16546)
; OTHER INFORMATION: n is 'g' or 'a'
; US-11-005-030-1

Query Match      28.8%; Score 502.8; DB 15; Length 26323;
Best Local Similarity 87.1%; Pred. No. 1.3e-148;
Matches 552; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1110 AAATGATCATGACGAAGAAATTTGTGGTCAATGAGATCACTGTGCTTGTTCGAGC 1169
DB 24690 ACAGGCTTCGACGAAGAAATCGTTTGGTTATGCGAATCAACAGTGTGTTGTTGGAGC 24749
QY 1170 ATCTGCAACAGCAGGCTTTTGTCTGACGAAGCTGTGTATGGGCTTGTGTAAGCTC 1229
DB 24750 ATCTGCAACAGCAGGCTTTTGTCTGACGAAGCTGTGTATGGGCTTGTGTAAGCTC 24809
QY 1230 TGACCTGTCTACATCATCTTCCACAGAGCTGCTGTGTACTTCTTCAAAAGAAC 1289
DB 24810 TGACCTGTCTACATCATCTTCCACAGAGCTGCTGTGTACTTCTTCAAAAGAAC 24869
QY 1290 CAACACTATGAGGAGCTGTGCTGTATATTTTGGACTATTTCTTGAGAAATTAAGAGG 1349
DB 24870 CAACACTATGAGGAGCTGTGCTGTATATTTTGGACTATTTCTTGAGAAATTAAGAGG 24929
QY 1350 AGAGCCATATCTATCTTGAGAGCCCTTATCTTTTACCTGTGTATTAATCTGACAGAA 1409
DB 24930 AGAGCCATATCTATCTTGAGAGCCCTTATCTTTTACCTGTGTATTAATCTGAGATTA 24989
QY 1410 TGGTATATACATCAGAGGTTCCCAATTTAAACTCTCTCCATGTTACTCAATCTTTTAC 1469
|||||
```

```

DB 24990 TGGTATATATATACGAAGAAATTTCCATTTAAACACTTCCATGTTACTATCTTTAC 25049
QY 1470 CAACATTTGTGTTTCTTATCTAGCCAACTATCTATTTGAAAGTGAACCTTCCCAA 1529
DB 25050 CAACATTTGCATCTCTATCTTAGCCAACTATCTATTTGAAAGTGAACCTTCCCAA 25109
QY 1530 ATTAGATGATTTGATGCTGTTGTGCGAAGGACAGTGAAGAAACATGACAGACCAT 1589
DB 25110 ATTAGATGATTTGATGCTGTTGTGCGAAGGACAGTGAAGAAACATGATTAAGACAT 25169
QY 1590 TCTAGTCAGAAATGAAATATCAATTAATGAAGAACTTGCACCTGGAACCTCGGACAG 1649
DB 25170 TCTGTCAAAAATGAAATATTAATTTAGATGAAGAACTTGCACCTGGAAGCCACGACAG 25229
QY 1650 CCTAACCTCAGTTCACTTTACCAATTAAGAGGAGCCCTCTGATGTTGATTCAGTTC 1709
DB 25230 CATGACCTCAGCTCACTTACCAATTAAGAGGAGCCCTCTGATGTTGATTCAGTTC 25289
QY 1710 GAGAGGCTCTGGAGCTGAAGATTAATTAATGA 1743
DB 25290 AGAAGGCTCTGGAGCTGAAGATTAATTAATGA 25323

RESULT 12
US-09-911-077A-19
; Sequence 19, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:00805
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 19
; LENGTH: 119040
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (2347)..(90873)
; OTHER INFORMATION: N = A, C, G or T/U
; US-09-911-077A-19

Query Match      28.8%; Score 502.8; DB 3; Length 119040;
Best Local Similarity 87.1%; Pred. No. 3.5e-148;
Matches 552; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1110 AAATGATCATGACGAAGAAATTTGTGGTCAATGAGATCACTGTGCTTGTTCGAGC 1169
DB 30755 ACAGGCTTCGACGAAGAAATCGTTTGGTTATGCGAATCAACAGTGTGTTGTTGGAGC 30814
QY 1170 ATCTGCAACAGCAGGCTTTTGTCTGACGAAGCTGTGTATGGGCTTGTGTAAGCTC 1229
DB 30815 ATCTGCAACAGCAGGCTTTTGTCTGACGAAGCTGTGTATGGGCTTGTGTAAGCTC 30874
QY 1230 TGACCTGTCTACATCATCTTCCACAGAGCTGCTGTGTACTTCTTCAAAAGAAC 1289
DB 30875 TGACCTGTCTACATCATCTTCCACAGAGCTGCTGTGTACTTCTTCAAAAGAAC 30934
QY 1290 CAACACTATGAGGAGCTGTGCTGTATATTTTGGACTATTTCTTGAGAAATTAAGAGG 1349
DB 30935 CAACACTATGAGGAGCTGTGCTGTATATTTTGGACTATTTCTTGAGAAATTAAGAGG 30994
QY 1350 AGAGCCATATCTATCTTGAGAGCCCTTATCTTTTACCTGTGTATTAATCTGACAGAA 1409
DB 30995 AGAGCCATATCTATCTTGAGAGCCCTTATCTTTTACCTGTGTATTAATCTGAGATTA 31054
QY 1410 TGGTATATACATCAGAGGTTCCCAATTTAAACTCTCTCCATGTTACTCAATCTTTTAC 1469
|||||
```

Db 31055 TGGTATATATATACAGAAATTTCCATTAAACACTTGCCATGGTTACATCATTTCTTAAC 31114
QY 1470 CAACATTGTTGTTCTTATCTAGCAAGATATCTATTGGAAGTGAACCTTGCTCCAAA 1529
Db 31115 CAACATTGTTGTTCTTATCTAGCAAGATATCTATTGGAAGTGAACCTTGCTCCAAA 31174
QY 1530 ATTAGATGATTTGATGCTGTTGTGCGAAGGCAAGTGAAGAGAACATGACAAACAT 1589
Db 31175 ATTAGATGATTTGATGCTGTTGTGCGAAGGCAAGTGAAGAGAACATGACAAACAT 31234
QY 1590 TCTAGTCAGAAATGAAATATCAAAATTAATGAATCTGCACTGTTGAAACCTTGCGAGAG 1649
Db 31235 TCTGTCAAAAATGAAATATTAATTAATGAACTTGCACTGTTGAAAGCCACAGACAGAG 31294
QY 1650 CCTAACCCCTCAGTTCACTTGCACTTAAGAGAGGCCCTCTGATGTTGATTCAGCTCC 1709
Db 31295 CATGACCCCTCAGCTCAACTTTCACCAATTAAGAGGCCCTCTGATGTTGATTCAGCTCC 31354
QY 1710 GGAGGGGTCTGGGACTGAAGATTAATTACATGA 1743
Db 31355 AGAAGGGTCTGGGACTGAAGATTAATTACATGA 31388

RESULT 13

US-10-724-806-19
; Sequence 19, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARUNDARAM, SUBRAMANIAM
; APPLICANT: PERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/10/724,806
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; PRIOR FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 119040
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (2347)..(90873)
; OTHER INFORMATION: N = A, C, G or T/U
US-10-724-806-19

Query Match 28.8%; Score 502.8; DB 9; Length 119040;
Best Local Similarity 87.1%; Pred. No. 3.5e-148;
Matches 552; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1110 AATGATCAGACAGAAATTTGTGGTCATGAGATCATCTGCTTGTTCGAGC 1169
Db 30755 ACAGGCTTGGCAAGAAATCGTTGGTTATGCCAATCACAGTGTGTTGGAGC 30814
QY 1170 ATTCGACACGCGATGCTTGTGACGAAGACTGTGTATGGGCTCTGGTAACCTGAGCTC 1229
Db 30815 ATTCGACACGCGATGCTTGTGACGAAGACTGTGTATGGGCTCTGGTAACCTGAGTTC 30874
QY 1230 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGTACTCTCTCAAAAGAAC 1289
Db 30875 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGTACTCTCTCAAAAGAAC 30934
QY 1290 CAACCTTATGGGCGAGTTGCTGTTATATTTTGAATTAATCTCTGAAGATTAAGAGG 1349
Db 30935 CAACCTTATGGGCGAGTTGCTGTTATATTTTGAATTAATCTCTGAAGATTAAGAGG 30994
QY 1350 AGAGCCATATCTATCTTCTGAGCCCTTAATCTTACCCTGTTATTAATCTGACAAAGAA 1409
Db 30995 GAGCCATATCTATCTTCTGAGCCCTTAATCTTACCCTGTTATTAATCTGACAAAGAA 31054

QY 1410 TGGTATATATATACAGAGGTTCCATTAAACCTCTCCATGTTACCTCATTTCTTAC 1469
Db 31055 TGGTATATATATACAGAAATTTCCATTAAACACTTGCCATGGTTACATCATTTCTTAAC 31114
QY 1470 CAACATTGTTGTTCTTATCTAGCAAGATATCTATTGGAAGTGAACCTTGCTCCAAA 1529
Db 31115 CAACATTGTTGTTCTTATCTAGCAAGATATCTATTGGAAGTGAACCTTGCTCCAAA 31174
QY 1530 ATTAGATGATTTGATGCTGTTGTGCGAAGGCAAGTGAAGAGAACATGACAAACAT 1589
Db 31175 ATTAGATGATTTGATGCTGTTGTGCGAAGGCAAGTGAAGAGAACATGACAAACAT 31234
QY 1590 TCTAGTCAGAAATGAAATATCAAAATTAATGAATCTGCACTGTTGAAACCTTGCGAGAG 1649
Db 31235 TCTGTCAAAAATGAAATATTAATTAATGAACTTGCACTGTTGAAAGCCACAGACAGAG 31294
QY 1650 CCTAACCCCTCAGTTCACTTGCACTTAAGAGAGGCCCTCTGATGTTGATTCAGCTCC 1709
Db 31295 CATGACCCCTCAGCTCAACTTTCACCAATTAAGAGGCCCTCTGATGTTGATTCAGCTCC 31354
QY 1710 GGAGGGGTCTGGGACTGAAGATTAATTACATGA 1743
Db 31355 AGAAGGGTCTGGGACTGAAGATTAATTACATGA 31388

RESULT 14

US-09-911-077A-14
; Sequence 14, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARUNDARAM, SUBRAMANIAM
; APPLICANT: PERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 142299
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
; NAME/KEY: modified base
; LOCATION: (1305)..(127835)
; OTHER INFORMATION: N = A, C, G or T/U
US-09-911-077A-14

Query Match 28.8%; Score 502.8; DB 3; Length 142299;
Best Local Similarity 87.1%; Pred. No. 4e-148;
Matches 552; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1110 AATGATCAGACAGAAATTTGTGGTCATGAGATCATCTGCTTGTTCGAGC 1169
Db 94673 ACAGGCTTGGCAAGAAATCGTTGGTTATGCCAATCACAGTGTGTTGGAGC 94732
QY 1170 ATTCGACACGCGATGCTTGTGACGAAGACTGTGTATGGGCTCTGGTAACCTGAGCTC 1229
Db 94733 ATTCGACACGCGATGCTTGTGACGAAGACTGTGTATGGGCTCTGGTAACCTGAGTTC 94792
QY 1230 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGTACTCTCTCAAAAGAAC 1289
Db 94793 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGTACTCTCTCAAAAGAAC 94852
QY 1290 CAACCTTATGGGCGAGTTGCTGTTATATTTTGAATTAATCTCTGAAGATTAAGAGG 1349
Db 94853 CAACCTTATGGGCGAGTTGCTGTTATATTTTGAATTAATCTCTGAAGATTAAGAGG 94912
QY 1350 AGAGCCATATCTATCTTCTGAGCCCTTAATCTTACCCTGTTATTAATCTGACAAAGAA 1409

Db	94913	GGAGCCATATCTGATCTTACGCCCTTGATCTTACCCGGGATATACCGTATACCGTATGATATA	949722
Qy	1410	TGGTATATACAAATCAGAGGTTCCCATTTAAAACCTCTCCATGGTTACCTCATTTCTTTAC	1469
Db	94973	TGGTATATATAATCAGAAATTTCCATTTAAAACACTGGCATGGTATACATTTCTTAAAC	950322
Qy	1470	CAACATTTGTGTTTCTTATCTTACGCCAAGTATCTATTTGAAAGTGAACCTTGCTCCAAA	1529
Db	95033	CAACATTTGCATCTCCATCTTACGCCAAGTATCTATTTGAAAGTGAACCTTGCCACCTTAA	950922
Qy	1530	ATTAGATGTATTTTGATGCTGTTGTCGCAGGCCACAGTGAAGAGACATGACACACAT	1589
Db	95093	ATTAGATGTATTTTGATGCTGTTGTCGCAGGCCACAGTGAAGAAACATGATTAAGACATAT	951522
Qy	1590	TCTATGAGAAATGAATAATATCAAAATTAATGSAACCTTGCAACCTCGGAGAG	1649
Db	95153	TCTTGTCAAAAATGAATAATATTAATTAATGATGSAACCTTGCACTGTATAGCCACGACAGAG	952122
Qy	1650	CCTAACCCCTCAGTTCAACTTTCACCAATAAGAGGCCCTCTCTGATGTGATTTCCAGTCC	1709
Db	95213	CATGACCCCTCAGCTCAACTTTCACCAATAAGAGGCCCTCTCTGATGTGATTTCCAGTCC	952722
Qy	1710	GGAGGGGCTCTGGACCTGAAGATATTTTCAATGA	1743
Db	95273	AGAAGGGTCTGGACCTGAAGATATATTTTCAAGTGA	95306

```

RESULT 15
US-10-724-806-14
; Sequence 14, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARASUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAM
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT.008US
; CURRENT APPLICATION NUMBER: US/10/724,806
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; PRIOR FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 142299
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (1305)..(127835)
; OTHER INFORMATION: N = A, C, G or T/U
; US-10-724-806-14

```

Query Match	28.8%;	Score 502.8;	DB 9;	Length 142239;
Best Local Similarity	87.1%;	Pred. No. 4e-148;		
Matches 552;	Conservative 0;	Mismatches 82;	Indels 0;	Gaps 0;

Accession	Sequence	Position
QY	AAATGATTCAGACAAGAAATTGTGTGGCCATGAGATCATCTGTCTGTGTGGAC	1169
Db	ACAGGCTTCGGACAAAGAAATCGTTTGGGTTATGGCAATCACAGTCTTGTCTTGGAC	947323
QY	ATCTGCAACAGCCATGGCTTGTCTGAGAGAACTGTGTATGGGCTCTGTACTGAGCTC	1229
Db	ATCTGCAACAGCCATGGCTTGTCTGAGAGAACTGTGTATGGGCTCTGTGACTCACTTC	947922
QY	ATCTGCAACAGCCATGGCTTGTCTGAGAGAACTGTGTATGGGCTCTGTACTGAGCTC	1289
Db	ATCTGCAACAGCCATGGCTTGTCTGAGAGAACTGTGTATGGGCTCTGTGACTCACTTC	948522

QY	1290	AAACACTAATGGGGGAGGTGCTGGGTATATATTTTGGACATATTCGAGAAATTAACGAGAG	1349
Db	94853	CAACACTAATGGGGGCGGTGGCAGAGTTATGTTCTGCGCTTTCGAGAAATTAACGAGAG	94912
QY	1350	AGAGCCATATCTATCTACTTGGAGCCCTTAAATCTTACCCCTGGTATTACTCGACAAAGAA	1409
Db	94913	GGAGCCATATCTATCTATCTACTTGGAGCCCTTAAATCTTACCCCTGGTATTACTCGATGA	94972
QY	1410	TGGTATATTAACATCAGAGGTGCTCCATTTAAATCTCTCCATGGTATACCTCATTCCTTAC	1469
Db	94973	TGGTATATTAACATCAGAAATTTTCATTTAAACCTTGCGATGGTTACATCATTTCTTAAC	95032
QY	1470	CAACATTTGTGTTTCTTATCTATCCCAAGATCATTTTGAAGTGAACCTTGCTCCAAA	1529
Db	95033	CAACATTTGGATCTCCCATCTACCAAGATATCATTTGAAGTGAACCTTGCCACCTTAA	95092
QY	1530	ATTAGATGTAATTTGATGCTGTGTCCGAAGGCACTGAAGAAACATGGAACAAGCCAT	1589
Db	95093	ATTAGATGTAATTTGATGCTGTGTCCGAAGGCACTGAAGAAACATGGAATGAACAT	95152
QY	1590	TCTAGTCAGAAATGAATAATTCAAATTAATGAACCTTGACCTGTGAACCTCGGAGAG	1649
Db	95153	TCTGTGTAATAAATGAATAATTTTAAATTAATGAACCTTGACCTGTGTAACCCACGACAG	95212
QY	1650	CTTAAACCTCAGTTCAAATTTCCACAAATAGAAGAGCCCTCTTGATGTTGATTCAGTCC	1709
Db	95213	CATGACCTCAGCTCAACCTTCCACCAATGAAGAGGCCCTCTTGATGTTGATTCAGTCC	95272
QY	1710	GGAGGGGTCTGGAGCTGAAGATTAATTAACATTA	1743
Db	95273	AGAAGGGTCTGGAGCTGAAGATTAATTAACGTTA	95306

Search completed: July 10, 2006, 14:50:57
Job time : 2000 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:47 ; Search time 258 seconds
(without alignments)
8783.456 Million cell updates/sec

Title: US-10-724-806-3

Perfect score: 1743
Sequence: 1 atgccttcacatgctgaag.....ctgaagataattacaarga 1743

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37.6	2.2	933	US-11-266-748A-359132	Sequence 359132,
2	37.6	2.2	933	US-11-266-748A-386971	Sequence 386971,
3	37.6	2.2	933	US-11-266-748A-442511	Sequence 442511,
4	37.6	2.2	1705	US-10-449-902-9396	Sequence 9396, Ap
5	37	2.1	518	US-11-266-748A-14121	Sequence 14121, A
6	37	2.1	628	US-11-266-748A-75909	Sequence 75909, A
7	37	2.1	628	US-11-266-748A-109134	Sequence 109134,
8	37	2.1	628	US-11-266-748A-128720	Sequence 128720,
9	37	2.1	650	US-11-266-748A-48029	Sequence 48029, A
10	37	2.1	801	US-11-266-748A-10950	Sequence 10950, A
11	37	2.1	874	US-11-266-748A-177977	Sequence 177977,
12	37	2.1	902	US-11-266-748A-114600	Sequence 114600,
13	37	2.1	902	US-11-266-748A-130654	Sequence 130654,
14	37	2.1	1000	US-11-266-748A-119114	Sequence 119114,
15	37	2.1	1000	US-11-266-748A-161278	Sequence 161278,
16	37	2.1	1124	US-11-266-748A-178110	Sequence 178110,
17	37	2.1	1464	US-11-266-748A-227710	Sequence 227710,
18	37	2.1	1681	US-11-266-748A-84832	Sequence 84832, A
19	37	2.1	1681	US-11-266-748A-111914	Sequence 111914,
20	37	2.1	1681	US-11-266-748A-137643	Sequence 137643,
21	37	2.1	3214	US-11-266-748A-278774	Sequence 278774, A
22	36.6	2.1	1261	US-11-266-545-342	Sequence 342, App
23	36.6	2.1	165280	US-11-021-837-47	Sequence 47, Appl
24	36	2.1	1768	US-11-266-748A-73373	Sequence 73373, A

25	36	2.1	1768	US-11-266-748A-107799	Sequence 107799,
26	36	2.1	1768	US-11-266-748A-126184	Sequence 126184,
27	35.8	2.1	755217	US-11-266-748A-39045	Sequence 29045, A
28	35.4	2.0	1791	US-10-449-902-4799	Sequence 4799, Ap
29	35.4	2.0	2549	US-10-449-902-16653	Sequence 16653, A
30	35	2.0	804	US-11-266-748A-86191	Sequence 86191, A
31	35	2.0	804	US-11-266-748A-139002	Sequence 139002,
32	35	2.0	1000	US-11-266-748A-199625	Sequence 199625,
33	35	2.0	42790	US-11-266-748A-23980	Sequence 23980, A
34	34.8	2.0	100096	US-11-266-748A-59715	Sequence 59715, A
35	34.2	2.0	1099	US-11-266-748A-164467	Sequence 164467,
36	34	2.0	1000	US-11-266-748A-199378	Sequence 199378,
37	34	2.0	1962	US-11-266-748A-32112	Sequence 32112, A
38	34	2.0	2207	US-11-266-748A-25770	Sequence 25770, A
39	34	2.0	2876	US-10-505-928-468	Sequence 468, App
40	34	2.0	2876	US-11-341-567-1	Sequence 1, Appl
41	34	2.0	5262	US-11-217-525-76367	Sequence 76367, A
42	34	2.0	9360	US-11-266-748A-28672	Sequence 28672, A
43	34	2.0	14544	US-11-266-748A-23683	Sequence 23683, A
44	34	2.0	17509	US-11-266-748A-28800	Sequence 28800, A
45	33.8	1.9	397	US-11-266-748A-37469	Sequence 37469, A

ALIGNMENTS

RESULT 1
US-11-266-748A-359132
; Sequence 359132, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Johnston, Paul
; APPLICANT: Harkin, Paul
; APPLICANT: Molligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 359132
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-11-266-748A-359132

Query Match 2.2%; Score 37.6; DB 8; Length 933;
Best Local Similarity 45.1%; Prod. No. 0.4; 264; Indels 6; Gaps 2;
Matches 222; Conservative 0; Mismatches

QY 119 AACCCATCATAGTGGGGCCCGTACATGTTGTTGTTGTTGTTTACCATAGAG 178
DB 279 AAGGCTACTTCCTCGCTGAGGGGACATGTTGTTGTTGTTGTTGTTT 338
QY 179 CCACCTGGTTGAGAGAGCTACATGATGAGACAGAGAGTATGGCCAGGTT 238

Db 339 CCAGCAATGTTGGAGTGGACATTTTCATGCGCTGGCAGGGCTCAGGTCGTCTACGGGCA 398
Qy 239 GTGGCTAGCTTGGGCTGATGACCAATGATATTTCTGTAGTCTAATTTTAGGTGTC 298
Db 339 TTTCTGTA---TCAGCTATGAACTTAATGCTTTTCTGTGTGTAATGTTGGCTGGA 455
Qy 239 TGTTTTTCGCAAACTTATGCTTCCAAAGGATATGTGATCTATTAGACCATTCGAAAC 358
Db 456 TCTTCTTACCCATCTACATGTCGTGTCAGGTCACCAAGATGCAAGATATCTACGGAAGC 515
Qy 359 AGATCTATGAAAGGCAATGGGTGGGCTGCTTTCATCCCTGACATGATGGAGAGATGT 418
Db 516 GCTTGGTGGCATGCAATCCCATCATCTGCTGCTGATCTTACCTATTATCTACATCT 575
Qy 419 T---CTGGGCTGACCAATTTTCTGTGATTTAGGGGCCACATCAGCCGTATCATTTAG 475
Db 576 TCACCAAGATCTCGGTAGACATGTATGAGGTGCCATCTTATCCAGAGCTTTGGACAC 635
Qy 476 TGGATGTGAACATATCGGTCATGTCTCTGCACTGATTTGCCATTTCTTATACCTTAGTG 535
Db 636 TGGATCTGTACTGTGCAATGTTGGGCTAATGCGCATCATCTGCTATATACAGGTTGCTG 695
Qy 536 GTGGGCTCTACTCTGTGGCATATACGATGTTGTCAGCTATTTCTGCAATTTTATAGAGC 595
Db 696 GTGGCTGGCTGCTGTGATCTACAGGATGCCCTGCAAGCGCTGATCATGCTTATAGAG 755
Qy 596 TGTGATCAGTG 607
Db 756 CGCTCACCCTTG 767

RESULT 2

US-11-266-748A-386971
; Sequence 386971, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 386971
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-386971

Query Match 2.2%; Score 37.6; DB 8; Length 933;
Best Local Similarity 45.1%; Pred. No. 0.4;
Matches 222; Conservative 0; Mismatches 264; Indels 6; Gaps 2;
Qy 119 AAGCATATAGTCCGGGCGCTGACATTTGTTGTGTGTGTGTTTACATGACAG 178

Db 279 AAGCTACTTCTGCTGAGGAGGACATGATGTGTGTCAGGTGGTGCATCTTTGTTTG 338
Qy 179 CCACCTGGGTTGGAGAGGCTTACATCATATGGGAGACAGAGACAGTATGGGACAGTT 238
Db 339 CCAGCAATGTTGGAGTGGACATTTTATGCTTGGCTGGCAGGGTCAAGTGTCTGTAAGGGA 398
Qy 239 GTGGCTAGCTTGGGCTGATGACCAATGATATTTCTGTAGTCTAATTTTAGGTGTC 298
Db 339 TTTCTGTA---TCAGCTATGAACTTAATGCTTTTCTGTGTGTAATGTTGGCTGGA 455
Qy 299 TGTTTTTCGCAAACTTATGCTTCCAAAGGATATGTGATCTATTAGACCATTCGAAAC 358
Db 456 TCTTCTTACCCATCTACATGTCGTGTCAGGTCACCAAGATGCAAGATATCTACGGAAGC 515
Qy 359 AGATCTATGAAAGGCAATGGGTGGGCTGCTTTCATCCCTGACATGATGGAGAGATGT 418
Db 516 GCTTGGTGGCATGCAATCCCATCATCTGCTGCTGATCTTACCTATTATCTACATCT 575
Qy 419 T---CTGGGCTGACCAATTTTCTGTGATTTAGGGGCCACATCAGCCGTATCATTTAG 475
Db 576 TCACCAAGATCTCGGTAGACATGTATGAGGTGCCATCTTATCCAGAGCTTTGGACAC 635
Qy 476 TGGATGTGAACATATCGGTCATGTCTCTGCACTGATTTGCCATTTCTTATACCTTAGTG 535
Db 636 TGGATCTGTACTGTGCAATGTTGGGCTAATGCGCATCATCTGCTATATACAGGTTGCTG 695
Qy 536 GTGGGCTCTACTCTGTGGCATATACGATGTTGTCAGCTATTTCTGCAATTTTATAGAGC 595
Db 696 GTGGCTGGCTGCTGTGATCTACAGGATGCCCTGCAAGCGCTGATCATGCTTATAGAG 755
Qy 596 TGTGATCAGTG 607
Db 756 CGCTCACCCTTG 767

RESULT 3

US-11-266-748A-442511/c
; Sequence 442511, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 442511
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-442511

Query Match 2.2%; Score 37.6; DB 8; Length 933;
 Best Local Similarity 45.1%; Pred. No. 0.4; Mismatches 264; Indels 6; Gaps 2;
 Matches 222; Conservative 0;

```

Oy 119 AAGCATCATAGTGGGGGCGGTGACATTTGTTGGTGGTGGTTTACCATGACAG 178
Db 655 AAGGCTACTTCTGCTGGAGGGGACATGTTGGTGGCCAGTGGGTGATCTTTGGT 596
Oy 179 CCACCTGGGTTGGAGAGGCTCATCATANGGACAGCAGAAACATGTAAGGCGCAGTT 238
Db 595 CCAGCAATGTTGGAAGTGGACATTTTCATTTGCGCTGGCAGGTGCTGCTACGGGCA 536
Oy 239 GTGTCTAGCTTGGGCTCATGACCCATTTGATTTCTGAGTCTAATTTTGGTGTGTC 298
Db 535 TTTCTGTA---TCAGTTATGAATTAATGAGCTTTGTTCTGTCGATGTTGGCTGGA 479
Oy 299 TGTTTTTCGGAACCTATGCGTTCCAGGGAATGTAATGTAATTTAGACCATTCACAC 358
Db 478 TCTTCTACCCATCTACATTTGCTGGTCAGTCCACAGATGCCAGAAATACCTACGGAAGC 419
Oy 359 AGATCTATGGAAGGCGATGGGCTGCTCTTATCCCTGCACCTGATGGAGATGT 418
Db 418 GCTTGGTGGCATCAGAAATCCCATCATCTGCTGCTACTTACCTATTATCTACATCT 359
Oy 419 T---CTGGGCTGACGCAATTTTCTCTGCAATTAAGGGCCACATCAGCGTATCATTTGATG 475
Db 358 TCACCAAGATCTCGGTAGACATGTATGACAGTGCATCTTTCACGACGCTTTGGCAC 299
Oy 476 TGGATGTGAACATATGCGTATTTCTCTGCACTCATTTCCATTTTATACCTAATGG 535
Db 298 TGGATCTGTAACCTGGCCATAGTTGGGCTACCTGGCCATCTGTATACACGGTTGGC 239
Oy 536 GTGGGCTCTACTCTGGGATATCTGATGTTGTCCAGTATTTTCGATTTTATAGAGC 595
Db 238 GTGGGCTGCTGCTGTGATCTACAGGATGCCCTGACAGCGTATCATGTTATAGAGG 179
Oy 596 TGTGATCAGTG 607
Db 178 CGCTACACTTTG 167

```

RESULT 4

US-10-449-902-9396
 ; Sequence 9396; Application US/10449902
 ; Publication No. US20060123505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: National Institute of Agricultural Sciences.
 ; APPLICANT: Bio-oriented Technology Research Advancement Institution.
 ; APPLICANT: The Institute of Physical and Chemical Research.
 ; APPLICANT: Foundation for Advancement of International Science.
 ; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
 ; FILE REFERENCE: MOA-A0205Y1-US
 ; CURRENT APPLICATION NUMBER: US/10/449, 902
 ; CURRENT FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: JP 2002-203269
 ; PRIOR FILING DATE: 2002-05-30
 ; PRIOR APPLICATION NUMBER: JP 2002-383870
 ; PRIOR FILING DATE: 2002-12-11
 ; NUMBER OF SEQ ID NOS: 56791
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 9396
 ; LENGTH: 1705
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: AK106744
 ; DATABASE ENTRY DATE: 2002-08-28
 ; US-10-449-902-9396

Query Match 2.2%; Score 37.6; DB 6; Length 1705;
 Best Local Similarity 57.8%; Pred. No. 0.58; Mismatches 49; Indels 0; Gaps 0;
 Matches 67; Conservative 0;

```

Oy 1276 TTCATCAAGGACCAACACTTATGGGCGAGTGGCTGTATATTTGACTATTCCTG 1335
Db 413 TTCATTCAGCAACATTTGCTGTGGTGCAGTCTGTGATCTTTGGTGAATTAAGTG 472
Oy 1336 AGAATTAAGGAGGAGGACCATATCTATCTTACGAGCCCTTAATCTTACCTCTG 1391
Db 473 GCATTGATTTGTAAGGCGCCCTTATTCATTAATCTTCAAGTTACCATCATCAATTGG 528

```

RESULT 5

US-11-266-748A-14121
 ; Sequence 14121; Application US/11266748A
 ; Publication No. US20060134663A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harkin, Paul
 ; APPLICANT: Mulligan, Patrick
 ; TITLE OF INVENTION: Transcriptome Microarray Technology and
 ; TITLE OF INVENTION: Methods of Using the Same
 ; FILE REFERENCE: 55815-0102 (319189)
 ; CURRENT APPLICATION NUMBER: US/11/266,748A
 ; CURRENT FILING DATE: 2005-11-03
 ; PRIOR APPLICATION NUMBER: EP 04105479.2
 ; PRIOR FILING DATE: 2004-11-03
 ; PRIOR APPLICATION NUMBER: EP 04105482.6
 ; PRIOR FILING DATE: 2004-11-03
 ; PRIOR APPLICATION NUMBER: EP 04105483.4
 ; PRIOR FILING DATE: 2004-11-03
 ; PRIOR APPLICATION NUMBER: EP 04105507.0
 ; PRIOR FILING DATE: 2004-11-03
 ; PRIOR APPLICATION NUMBER: EP 04105485.9
 ; PRIOR FILING DATE: 2004-11-03
 ; PRIOR APPLICATION NUMBER: EP 04105484.2
 ; PRIOR FILING DATE: 2004-11-03
 ; PRIOR APPLICATION NUMBER: US 60/662,276
 ; PRIOR FILING DATE: 2005-03-14
 ; PRIOR APPLICATION NUMBER: US 60/700,293
 ; PRIOR FILING DATE: 2005-07-18
 ; NUMBER OF SEQ ID NOS: 483996
 ; SOFTWARE: Patentin version 3.3
 ; SEQ ID NO 14121
 ; LENGTH: 518
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; US-11-266-748A-14121

Query Match 2.1%; Score 37; DB 8; Length 518;
 Best Local Similarity 64.7%; Pred. No. 0.42; Mismatches 30; Indels 0; Gaps 0;
 Matches 55; Conservative 0;

```

Oy 185 GGGTTGGAGGAGGACATCATATGGGACAGAGAGAGGTATGGGCGCGGTGTGTC 244
Db 371 GAGTTGGCCGAGAGATGATGATTTGGGACAGAAAGCCGTTCAGACCCACGTTGTCTC 430
Oy 245 TAGCTTGGGCTCATGACCCCATTTG 269
Db 431 TAGGCAAGGACCTTTGGCCCTTTG 455

```

RESULT 6

US-11-266-748A-75909
 ; Sequence 75909; Application US/11266748A
 ; Publication No. US20060134663A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harkin, Paul
 ; APPLICANT: Johnston, Patrick
 ; APPLICANT: Mulligan, Karl
 ; TITLE OF INVENTION: Transcriptome Microarray Technology and
 ; TITLE OF INVENTION: Methods of Using the Same
 ; FILE REFERENCE: 55815-0102 (319189)
 ; CURRENT APPLICATION NUMBER: US/11/266,748A
 ; CURRENT FILING DATE: 2005-11-03
 ; PRIOR APPLICATION NUMBER: EP 04105479.2

```

; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 75909
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-75909
```

```
Query Match      2.1%; Score 37; DB 8; Length 628;
Best Local Similarity 64.7%; Pred. No. 0.48;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
```

```
Oy 185 GGGTTGAGAGGCTACATCAATGAGACAGACAGTGTATGCGCCAGGTGTGCTC 244
Db 484 GAGTTGGCGGAGATGATGTGGGACGAGAAAGCCCTTGACCCCACTGTGTCTC 543
Oy 245 TAGCTTGGGCTCATGCACCCATTGG 269
Db 544 TAGGACGGGACCTTTGGCCCTTTG 568
```

```
RESULT 7
US-11-266-748A-109134
; Sequence 109134, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 109134
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-109134
```

```
Query Match      2.1%; Score 37; DB 8; Length 628;
Best Local Similarity 64.7%; Pred. No. 0.48;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
```

```
Oy 185 GGGTTGAGAGGCTACATCAATGAGACAGACAGTGTATGCGCCAGGTGTGCTC 244
Db 484 GAGTTGGCGGAGATGATGTGGGACGAGAAAGCCCTTGACCCCACTGTGTCTC 543
Oy 245 TAGCTTGGGCTCATGCACCCATTGG 269
Db 544 TAGGACGGGACCTTTGGCCCTTTG 568
```

```
RESULT 8
US-11-266-748A-128720/c
; Sequence 128720, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 128720
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-128720
```

```
Query Match      2.1%; Score 37; DB 8; Length 628;
Best Local Similarity 64.7%; Pred. No. 0.48;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
```

```
Oy 185 GGGTTGAGAGGCTACATCAATGAGACAGACAGTGTATGCGCCAGGTGTGCTC 244
Db 145 GAGTTGGCGGAGATGATGTGGGACGAGAAAGCCCTTGACCCCACTGTGTCTC 86
Oy 245 TAGCTTGGGCTCATGCACCCATTGG 269
Db 85 TAGGACGGGACCTTTGGCCCTTTG 61
```

```
RESULT 9
US-11-266-748A-48029
; Sequence 48029, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
```



```
/ Sequence 114600, Application US/11266748A
/ Publication No. US2006013463A1
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: Patentin version 3.3
/ SEQ ID NO: 114600
/ LENGTH: 902
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-114600

Query Match      2.1%; Score 37; DB 8; Length 902;
Best Local Similarity 64.7%; Pred. No. 0.6;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 1177 ACAGCATTGCTTCTTCGACGAGACTGTATGGGCTGTGACCTGACCTT 1236
    |||||
DB 523 AGAGCTTCGCTTCCAGTCGAGGATTATTAGCGGCTTCTTCTTACGCTC 464
    |||||

OY 1237 GTTACATCATCATCTTCCACAGC 1261
    |||||
DB 463 CATTTCCGATCATGATTCTCAGC 439
    |||||

RESULT 13
US-11-266-748A-120654
/ Sequence 120654, Application US/11266748A
/ Publication No. US2006013463A1
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
```

```
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: Patentin version 3.3
/ SEQ ID NO: 120654
/ LENGTH: 902
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-120654

Query Match      2.1%; Score 37; DB 8; Length 902;
Best Local Similarity 64.7%; Pred. No. 0.6;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 1177 ACAGCATTGCTTCTTCGACGAGACTGTATGGGCTGTGACCTGACCTT 1236
    |||||
DB 380 AGAGCTTCGCTTCCAGTCGAGGATTATTAGCGGCTTCTTCTTACGCTC 439
    |||||

OY 1237 GTTACATCATCATCTTCCACAGC 1261
    |||||
DB 440 CATTTCCGATCATGATTCTCAGC 464
    |||||

RESULT 14
US-11-266-748A-119114
/ Sequence 119114, Application US/11266748A
/ Publication No. US2006013463A1
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: Patentin version 3.3
/ SEQ ID NO: 119114
/ LENGTH: 1000
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-119114

Query Match      2.1%; Score 37; DB 8; Length 1000;
Best Local Similarity 64.7%; Pred. No. 0.64;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 185 GGGTGGAGAGGAGCTACATCATGAGGACAGAGAGGAGGAGGAGGAGGAGG 244
    |||||
DB 856 GAGTGGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 915
    |||||

OY 245 TAGCTTGGGCTCATGACCACTTGG 269
    |||||
DB 916 TAGGAGGAGGAGCTTGGGCCCTTTG 940
    |||||
```

Tue Jul 11 13:42:50 2006

us-10-724-806-3.rnpbn

Page 7

RESULT 15

US-11-266-748A-161278/c
; Sequence 161278, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 161278
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-161278

Query Match 2.1%; Score 37; DB 8; Length 1000;
Best Local Similarity 64.7%; Pred. No. 0.64;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 185 GGGTTGGAGAGGCTCATGATGGAGCAGAGCAGTATGGGCCAGTTGTGTC 244
Db 145 GAGTTGGCCGAGAGATGATGATTTGTGGCAGAGAGCCGTTGCAGCCCGTTGTGTC 86
QY 245 TAGCTTGGGCTCATGACCCCATTTGG 269
Db 85 TAGGCAGGAGCCTTTGGCCCTTTG 61

Search completed: July 10, 2006, 23:23:51
Job time : 261 secs

This Page Blank (uspio)

Db 241 GGCCTAGCTGGGCTCAGGCAACAAATTGATATTTCTTAGTCTGATTTTAGCTGACCTG 300
QY TTTTGTGGAACCTATGCGTTCGAAGGATATGTACTATGTAGACCATTTCAACAG 360
Db TTTCTTGGAAAACCTATGCGTTCGAAGGATATGTAGACCATTTTAGACCATTTCAACAG 360
QY ATCTATGGAAGGATATGCGTTCCTATCTCTGACCTGATGGAAGATGTTT 420
Db ATCTATGGAAGGATATGCGTTCCTATCTCTGACCTGATGGAAGATGTTT 420
QY TGGGCTGAGCAATTTTCTCTGCTTTGGAGCCACCATGAGGATCATGATGAT 480
Db TGGGCTGAGCAATTTTCTCTGCTTTGGAGCCACCATGAGGATCATGATGAT 480
QY GTGAACATATCGGTATGCTCTGACCTCATGTCATTTTATACCTTATGCGTGG 540
Db ATGACATATTCGTATCATCTCTGACCTCATGTCATTTTATACCTTATGCGTGG 540
QY CTCTATCTCTGAGCAATATCTGATGTTGCGAGCTATTCGATTTTATAGACTG 600
Db CTCTATCTCTGAGCAATATCTGATGTTGCGAGCTATTCGATTTTATAGACTG 600
QY CTCTATCTCTGAGCAATATCTGATGTTGCGAGCTATTCGATTTTATAGACTG 600
Db CTCTATCTCTGAGCAATATCTGATGTTGCGAGCTATTCGATTTTATAGACTG 600
QY ATCACTGCTCTTTTGGCTCTGACATCTGAGATCAACGATCGGATTCAGCTGT 660
Db ATCACTGCTCTCTTTGAGCTCTGACATCTGAGATCAACGATCGGATTCAGCTGT 660
QY CATGCTAAATACAGAGTCCCTGCTGAGAACATTTGAATCAGTTGAAGCTTACCT 720
Db CATGCTAAATACAGAGTCCCTGCTGAGAACATTTGAATCAGTTGAAGCTTACCT 720
QY CTGATATATTTCTGATATGATGCTGGTGAATCCCATGAGCAACCTTCCAGAG 780
Db CTGATATATTTCTGATATGATGCTGGTGAATCCCATGAGCAACCTTCCAGAG 780
QY GTCTCTCTCTTCACTCTGACCACTATGCTCAGATCTCTCTCTGAGCTTTTGG 840
Db GTCTCTCTCTTCTCTGACCACTATGCTCAGATCTCTCTCTGAGCTTTTGG 840
QY TGCCTGTGATGCTCTACCCGCAATGATAGAGCTATTTGAGCTTCAAGACTGG 900
Db TGCCTGTGATGCTCTACCCGCAATGATAGAGCTATTTGAGCTTCAAGACTGG 900
QY AACGAGACTGCTTACGGGATCCAGATCCAGAGCAAGAGAGAGAGAGATGATTC 960
Db AACGAGACTGCTTACGGGATCCAGATCCAGAGCAAGAGAGAGAGAGATGATTC 960
QY AACGAGACTGCTTACGGGATCCAGATCCAGAGCAAGAGAGAGAGATGATTT 960
Db AACGAGACTGCTTACGGGATCCAGATCCAGAGCAAGAGAGAGAGATGATTT 960
QY CGATGCTTCTGACATCTCTGCTGCTGATCATCTCTCTGAGCTTGTGCTGTT 1020
Db CGATGCTTCTGACATCTCTGCTGCTGATCATCTCTCTGAGCTTGTGCTGTT 1020
QY CCAATGTTTCTGACATCTCTGCTGCTGATCATCTCTCTGAGCTTGTGCTGTT 1020
Db CCAATGTTTCTGACATCTCTGCTGCTGATCATCTCTCTGAGCTTGTGCTGTT 1020
QY TCACTGCTGCTGATCTCTGACCTGATCTGTCATCTCTGCTGAGCTTGTGCT 1080
Db TCACTGCTGCTGATCTCTGACCTGATCTGTCATCTCTGCTGAGCTTGTGCT 1080
QY TCTGCTGCTGATCTCTGACCTGATCTGTCATCTCTGCTGAGCTTGTGCTG 1080
Db TCTGCTGCTGATCTCTGACCTGATCTGTCATCTCTGCTGAGCTTGTGCTG 1080
QY CGGAATATCTACCACTTCTCTGACAAATGATGATGAGAGAGAGAGAGAGAG 1140
Db CGGAATATCTACCACTTCTCTGACAAATGATGATGAGAGAGAGAGAGAGAG 1140
QY ATGAGATCATCTGCTGCTGCTGCTGAGATCTGCAACAGCAGCTTGTGCTGAG 1200
Db ATGAGATCATCTGCTGCTGCTGCTGAGATCTGCAACAGCAGCTTGTGCTGAG 1200
QY ACTGCTGATGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db ACTGCTGATGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
QY CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY TTTGAGATATCTCTGAGATTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 1380
Db TTTGAGATATCTCTGAGATTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 1380
QY TCTGCTGCTCTCTCTGAGATTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 1380
Db TCTGCTGCTCTCTCTGAGATTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 1380

QY 1381 TTCTACCTGCTTATTAATCTGACAAAGATGATATACAAATCAGAGGTTCCATTAA 1440
Db 1381 TTCTACCTGCTTATTAATTAATCTGATGATATGATATTAATTAATCAAGAAATTCATTAA 1440
QY 1441 ACTCTCTCAATGCTTACCTGATCTTATTTTCAAAATTTGTTTCTTATCTAGCCAGAT 1500
Db 1441 ACATCTGCTCAATGCTTACCTGATCTTATTTTCAAAATTTGTTTCTTATCTAGCCAGAT 1500
QY 1501 CTATTTGAAAGTGAACCTTCTCCAAATTTAGATGATTTGATGCTGTTGTCGAG 1560
Db 1501 CTATTTGAAAGTGAACCTTCTCCAAATTTAGATGATTTGATGCTGTTGTCGAG 1560
QY 1561 CACAGTGAAGAGACATGAGACCAATCTTACTGACAAATGAAATATCAATTAAT 1620
Db 1561 CACAGTGAAGAGACATGAGACCAATCTTACTGACAAATGAAATATCAATTAAT 1620
QY 1621 GAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAAC 1680
Db 1621 GAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAAC 1680
QY 1681 GAGGCTCTCTTGAATGATTTCCAGTCCGAGAGGCTTGGGACTGAAGATTAATCA 1740
Db 1681 GAGGCTCTCTTGAATGATTTCCAGTCCGAGAGGCTTGGGACTGAAGATTAATCA 1740
QY 1741 TGA 1743
Db 1741 TGA 1743

RESULT 2
US-09-974-300-501
; Sequence 501, Application US/0974300
; Patent No. 7018794
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09-974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 501
; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-501

Query Match 13.6%; Score 237.6; DB 5; Length 1461;
Best Local Similarity 51.9%; Pred. No. 3.3e-67;
Matches 670; Conservative 0; Mismatches 574; Indels 48; Gaps 4;

QY 124 ATCATATGCTGGGCGCTGACATTTGTTGTTGTTGTTTATCATGACAGCCACC 183
Db 98 ATGATCGTGGCAAAAAGAGATGCGTTTTCGTGGAATGATGATGATGCGGCA 157
QY 184 TGGGTTGAGAGAGCTATCAATGGAACAGAGAGAGAGAGAGAGAGAGAGAGAT 243
Db 158 TGGGTTGAGAGAGCTATCAATGGAACAGAGAGAGAGAGAGAGAGAGAGAGAT 211
QY 244 CTAGCTTGGGCTATGACCCATGAGATATCTGAGCTTAATTTTATGAGTCTGTT 303
Db 212 CTATCTGAGCCCAAGCCCTTGGGCTTACGATGATGAGCTGATATGAGCGGAT 271
QY 304 TTTGGAAGATATGCTTCAAGGATATGATGATGATGATGATGATGATGATGAT 363
Db 272 TTTGGAAGATATGCTTCAAGGATATGATGATGATGATGATGATGATGATGAT 331

```

QY 364 TATGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGAGAGATGTTCTGG 423
DB 332 TTCCGTAACCGGATGCGCGCGCTTTATATATACCGCGCTGTAGAGAAATGTTTGG 391
QY 424 GCTGCAGCAATTTCTCTGCATTTAGGGGCCACCATCAGCGTATCATTTGATGTGATG 483
DB 392 ACCGCGGCATCTTAACGGCATTTGGCAGCACTTTGCGAATGATTCGAAATATCATTTTC 451
QY 484 AACATATGGGTATGTCCTGACACATTTGCCATCTTTATATCCCTAGGGGTGGGCTC 543
DB 452 CAAACGTCATATATTTCTCGGATGATCGCATGCTTATACGGTGGCGGAGCATG 511
QY 544 TACTGTGGCAATATCTGATGTTGTCAGCTATTCGATTTTATAGACGTGGATC 603
DB 512 TGGGCGATGCTTTTCACAGATGCTTCAATATGATGCTATTTGCTGGGCTGTTTAT 571
QY 604 AGTGTCCCTTTTGCCCTGTCACATCCTGCACTCACGACATCGGATTCACAGCTGTG-- 660
DB 572 GTGTCGCCATTTGATTTGTAATGTCGATGCTCTGATAGCTATGGGCAATTTACAGG 631
QY 661 -----CATGCTAAATACCAGAGTCCCTGGCTGGGAAACCATTTGAT 700
DB 632 CACGATTTGCGCAGACGCGCAATCTGCTCGCGCGTTGACGCGTGGAAAAACCGGAT 691
QY 701 CAGTTGAA----GTCTACACCTGGCTTGATTTTCTGTATTGATGCTGGGTGGAATC 756
DB 692 TGGGGAACCTGTTTGGAACTGTGTGGATATGCGCTCCTCTTAATTTTGGAGGATATC 751
QY 757 CCATGGCAAGCTTACTTCAGAGGGTCTCTCTTCACTCTCAGCCACTATGCTCAGGTA 816
DB 752 GCATGGCAGGTGTACTTTAGCGCGTCTTTGCGCAAAATCGGAAAGCGCGCATGTGG 811
QY 817 CTGTCTCTCTGCGACCTTTGGGTGCGCTGTATGCTCTACCGCCATATGCAATAGA 876
DB 812 CAGTCATATATTTGCGGAGTATCTGATCATTTGCGGCAATTTGCGGCTGATCATTCGA 871
QY 877 GCTATTGAGACTTCCACAGACTGGAACAGACTGCTCAGGATATCCAGATCCCAAGACT 936
DB 872 GCTGCGGAAACAGTACCGATTTGAGCTGTTTGGAGCGACGCTCCGATTAACCGGGG 931
QY 937 AAGGAGAAAGCAGACATGATTTCTCCGATCGTTCTGCAAGTACCTGCGCTGTGTATC 996
DB 932 -----ATGATTTTGGCGCAAGCGTTGCGATTTGACGCGAGAAATCATC 976
QY 997 TCTTTCTTTGGGCTTGGTGTCTGTTTCACTGCTGTATGCTCTCAGCTGATCTGTATC 1056
DB 977 GCAGGCTCTGGGCTTGGGTGCAATGCGACGCGCTATGTCAAGCATGTGATCATGATTT 1036
QY 1057 CTGTGGCGAGTCTTATGTTTGTCTCGGAATATCTACAGCTTTCTTCAAGCAAAATGCA 1116
DB 1037 CTATCGGCATCATCAATGGCGCATGGAATATTTACCGCTCATCAAGCGGAAAGGCC 1096
QY 1117 TCAGACAGAGAAATTTGTGGGTCATGAGGATCATGTGCTTGTGTGGAGCATCTGCA 1176
DB 1097 ACCCAAAAACAGCTGCAAAAAGTCGCAACGCTCATATCTTTGTTGGCGGAGGAGCA 1156
QY 1177 AAGCAGCATGGCTTGTCTGACGAAGACTGTGTATGGGCTCTGTGATCCTGAGCTTGACTT 1236
DB 1157 GCGGTATGCGCGTGATGTCAAAAGCGTTTATATCTTATGATTTGGCTTGGAATTTA 1216
QY 1237 GTCTACATCATCATCTTCCACAGCTGCTGTGTACTCTTCAATCAAGAAACCAACACT 1296
DB 1217 GTTTATATGATCTTTTCCCGATTTACAAATGGCCCTCTTTTATTAAGAGAAATCTTT 1276
QY 1297 TATGGGCGAGTGTGCTGTTATATTTTGGACTATTTCTTGAAATTACTGAGAGAGCCCA 1356
DB 1277 TAGGGGTGATGTCTGATTTGCAATTTGCAATTTCTGAGCTCGCGGTGTGTGAACCC 1336
QY 1357 TATCTATCTGACGCCCTTAATCTTCAACC 1388
DB 1337 GCAATTGGGATTCGCGCGCTTGTGCGGATTC 1368

```

```

RESULT 3
US-09-270-767-1720
; Sequence 1720, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1720
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17002
RESULT 4
US-09-270-767-17002
; Sequence 17002, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 17002
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17002
Query Match 8.0%; Score 138.8; DB 3; Length 324;
Best Local Similarity 65.5%; Pred. No. 5,1e-35;
Matches 203; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
QY 808 GCTCAGTACTGTCTCTCTGCACTTTTGGGTGCTGTGATGCTTACCGCCATA 867
DB 15 GCCACGCTTCTGTCTATGTTGACGCGCGGATGATTTTGTATGATTCATTTCCCGGTG 74
QY 868 TGCATAGAGCTATTGGAGCTTCCACAGACTGGAACCAAGACTGCCATCGGATATCCAGAT 927
DB 75 CTCATCGAGAGCATATGCGCAAGGCTACACTTGGAAAGACATATTAACAGGACCTTAT 134
QY 928 CCAAGACTAAGAGAGACAGACATGATTTCTCCGATGTTCTGCAAGTACCTTGCCTC 987
DB 135 CCCCTACCGTGAACAGACAGACATGATTTCTGCCATGCTGTGATGATCTCAAGCCT 134
QY 988 GTGTACATCTCTCTTTGGGCTTGTGCTGTTTCACTGCTGTATGTTCTGACTGAC 1047
DB 195 GACTGTGTCTCTCTTTGATTTGGGCGCTGTTTCCGCGCGCTGATGTTCTCGCGGAC 254
QY 1048 TCGTCCATCTGTGCGGAGTTCTATGTTTGTCTGGAATATCAACAGCTTCTTCAAG 1107
DB 255 TCTCGTCTCTCTCCCGCGCTCCATGTTGCTCGGAACGTATGAATTTGATTTCCCT 314
QY 1108 CAAATGAT 1117
DB 315 CAGAGGCGT 324

```


REGISTRATION NUMBER: 34,287
REFERENCE/DOCKET NUMBER: 96700/393
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2839
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: OLIGONUCLEOTIDE
DESCRIPTION: NO
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: RAT
INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER
US-08-595-553A-1

Query Match 2.2%; Score 37.8; DB 3; Length 2839;
Best Local Similarity 49.3%; Pred. No. 0.63;
Matches 99; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 404 TGATGGAGAAGATGTTCTGGGCTGACAGCAATTTCTGCAATTAAGGGCCACCATCAGCG 463
DB 522 TGGTGGCCACAGTGTGATACAGGCATCGATCTACGCGCTGCTCATCTGAACC 581
QY 464 TGATCATGATGTGATGTAACAATATGGCATTTGTCTGCACTCATTCGACTTTCTT 523
DB 582 AAGTACCGGGTGGACATCTGGGCACTGCTCTGTCACAGAAATCATCTGACCTTGT 641
QY 524 ATACCTAGTGGGCTGCTACTCTGTCGATATATGATGTTGTCAGCTATTCTGCA 583
DB 642 ACACACCGTGGTGTGATGAAGCGCGTGTCTGACAGATGTGTTCCAGGTTGTGTTA 701
QY 584 TTTTATAGACTGTGATCA 604
DB 702 TGCTGTGGCTTCTGGGTGA 722

RESULT 11

US-09-995-007-1
Sequence 1, Application US/09995007
Patent No. 6803199
GENERAL INFORMATION:
APPLICANT: NANCY CARRASCO, ET AL.
TITLE OF INVENTION: THYROID SODIUM/IODIDE SYMPORTER AND
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
STREET: 90 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE
DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/995,007
FILING DATE: 26-NO. 6803199-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/595,553A
FILING DATE: FEBRUARY 1, 1996
ATTORNEY/AGENT INFORMATION:
NAME: CRAIG J. ARNOLD
REGISTRATION NUMBER: 34,287

REFERENCE/DOCKET NUMBER: 96700/393
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2839
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: <Unknown>
DESCRIPTION: OLIGONUCLEOTIDE
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-007-1

Query Match 2.2%; Score 37.8; DB 3; Length 2839;
Best Local Similarity 49.3%; Pred. No. 0.63;
Matches 99; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 404 TGATGGAGAAGATGTTCTGGGCTGACAGCAATTTCTGCAATTAAGGGCCACCATCAGCG 463
DB 522 TGGTGGCCACAGTGTGATACAGGCATCGATCTACGCGCTGCTCATCTGAACC 581
QY 464 TGATCATGATGTGATGTAACAATATGGCATTTGTCTGCACTCATTCGACTTTCTT 523
DB 582 AAGTACCGGGTGGACATCTGGGCACTGCTCTGTCACAGAAATCATCTGACCTTGT 641
QY 524 ATACCTAGTGGGCTGCTACTCTGTCGATATATGATGTTGTCAGCTATTCTGCA 583
DB 642 ACACACCGTGGTGTGATGAAGCGCGTGTCTGACAGATGTGTTCCAGGTTGTGTTA 701
QY 584 TTTTATAGACTGTGATCA 604
DB 702 TGCTGTGGCTTCTGGGTGA 722

RESULT 12

US-09-949-016-15637/C
Sequence 15637, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15637
LENGTH: 48794
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)...(48794)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15637

Query Match 2.2%; Score 37.8; DB 3; Length 48794;
Best Local Similarity 55.8%; Pred. No. 5.1;


```
Matches 72; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY 1414 ATATAATCAGAGGTTCCCATTTAAACCTCTCCATGGTATCCATCTTTTACCAC 1473
    |||||
Db 22938 ATATACATCTCTCTCTAGTTAAACCTATCTTTATCTGATCATCTTTAGACC 22879
    |||||
QY 1474 ATTTGTTCTTATCTAGCCAAATCTATTTGAAAGTGAACCTTGCCCAATTA 1533
    |||||
Db 22878 CTAATTCACAGATCTTTGCCATTTATGCTCTTAAACCTCGCCCTTACCTTACATTA 22819
    |||||
QY 1534 GATGATTT 1542
    |||||
Db 22818 TTGGTATCT 22810
    |||||

RESULT 13
US-09-513-999C-2240/c
/ Sequence 2240, Application US/09513999C
/ Patent No. 6783961
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Duclercq, A.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
/ Patent No. 6783961
/ FILE REFERENCE: 59 US2, REG
/ CURRENT APPLICATION NUMBER: US/09/513, 999C
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/122,487
/ NUMBER OF SEQ ID NOS: 36681
/ SOFTWARE: Patent.pm
/ SEQ ID NO 2240
/ LENGTH: 478
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 29..478
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 141
/ OTHER INFORMATION: y=c or t
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: 38
/ OTHER INFORMATION: Xaa=Leu or Pro
US-09-513-999C-2240

Query Match 2.1%; Score 37; DB 3; Length 478;
Best Local Similarity 64.7%; Pred. No. 0.31; 30; Indels 0; Gaps 0;
Matches 55; Conservative 0; Mismatches 30;

QY 1177 ACAGCCATGCTTGTGAGAGAGTGTATGGCTGTGACTGAGCTTGAACCTT 1236
    |||||
Db 91 AGAGCCTTCGCTTCCAGTCGAGGTTTATTATGAGCGCTCTTCTCTAGCTTGAGCTC 32
    |||||
QY 1237 GCTTACATATCATCTTCCACAGC 1261
    |||||
Db 31 CATTTCCGATCAGATTATCCAGC 7
    |||||

RESULT 14
US-09-539-333D-1/c
/ Sequence 1, Application US/09539333D
/ Patent No. 6476208
/ GENERAL INFORMATION:
/ APPLICANT: Cohen, Daniel
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Chumakov, Ilya
/ APPLICANT: Bougueleret, Lydie
/ APPLICANT: Bihain, Bernard
/ APPLICANT: Essioux, Laurent

/ TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
/ FILE REFERENCE: GENSET 047AUS
/ CURRENT APPLICATION NUMBER: US/09/539,333D
/ PRIOR FILING DATE: 2000-03-30
/ PRIOR APPLICATION NUMBER: US 60/126,903
/ PRIOR FILING DATE: 1999-03-30
/ PRIOR APPLICATION NUMBER: US 60/131,971
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: US 60/132,065
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: US 60/143,928
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: US 60/145,915
/ PRIOR FILING DATE: 1999-07-27
/ PRIOR APPLICATION NUMBER: US 60/146,453
/ PRIOR FILING DATE: 1999-07-29
/ PRIOR APPLICATION NUMBER: US 60/146,452
/ PRIOR FILING DATE: 1999-07-29
/ PRIOR APPLICATION NUMBER: US 60/162,288
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: US 09/416,384
/ PRIOR FILING DATE: 1999-10-12
/ NUMBER OF SEQ ID NOS: 231
/ SOFTWARE: Patent.pm
/ SEQ ID NO 1
/ LENGTH: 319608
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 31..1107
/ OTHER INFORMATION: 5'regulatory region g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 1108..1289
/ OTHER INFORMATION: exon A g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 14877..14920
/ OTHER INFORMATION: exon B g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 18778..18862
/ OTHER INFORMATION: exon Bbis g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 25593..25740
/ OTHER INFORMATION: exon C g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 29388..29502
/ OTHER INFORMATION: exon D g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 29967..30282
/ OTHER INFORMATION: exon E g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 64666..64812
/ OTHER INFORMATION: exon F g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 65505..65853
/ OTHER INFORMATION: exon G g35018 gene
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 65854..67854
/ OTHER INFORMATION: 3'regulatory region g35018 gene
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 94124..94964
/ OTHER INFORMATION: exon g35017
/ FEATURE:
```

NAME/KEY: exon
LOCATION: 201188..201234
OTHER INFORMATION: exon S g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 214676..214793
OTHER INFORMATION: exon T g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215702..215746
OTHER INFORMATION: exon U g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216836..216915
OTHER INFORMATION: exon V g35030 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 213818..215818
OTHER INFORMATION: 3' regulatory region g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215941
OTHER INFORMATION: exon R complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215975
OTHER INFORMATION: exon Rbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216661..216952
OTHER INFORMATION: exon Qbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 217027..217061
OTHER INFORMATION: exon Q complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 229647..229742
OTHER INFORMATION: exon X complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 230408..230721
OTHER INFORMATION: exon P complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231272..231412
OTHER INFORMATION: exon Obis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231787..231880
OTHER INFORMATION: exon O2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231870..231879
OTHER INFORMATION: exon O1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 2344174..2344321
OTHER INFORMATION: exon O complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 237406..237428
OTHER INFORMATION: exon Nbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239807
OTHER INFORMATION: exon N2 complement g34872 gene
FEATURE:
NAME/KEY: exon

LOCATION: 239719..239853
OTHER INFORMATION: exon N complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240569
OTHER INFORMATION: exon M1117 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240596
OTHER INFORMATION: exon M1090 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240617
OTHER INFORMATION: exon M1069 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240644
OTHER INFORMATION: exon M862 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240994
OTHER INFORMATION: exon M692 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..241685
OTHER INFORMATION: exon M1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240800..240993
OTHER INFORMATION: exon M51 complement g34872 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 241686..243685
OTHER INFORMATION: 5' regulatory region g34872 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 290652..292652
OTHER INFORMATION: 3' regulatory region g34665 gene
FEATURE:
NAME/KEY: exon
LOCATION: 292653..292841

Query Match 2.1%; Score 37; DB 3; Length 319608;
Best Local Similarity 47.3%; Pred. No. 37;
Matches 112; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 1418 ACAATCAAGGTTCCCATTTAAACTCTCTCCATGTTACCTCATTTCTTACCAACATTT 1477
DB 156436 ACATCTGATGTTCTCAGTCAAAACAAATTTCCACAAATCCCTAACCTGCTGATGAACGACA 156377
QY 1478 GTGTTCTTATCTACCAAGTATCTATTGAAAGTGAACCTTGCTCCAAATTAAGATG 1537
DB 156376 GAATTCATCTATTAACATGATGTCAGTTTCACCTGACGTAAGATTTTTCACATTTTG 156317
QY 1538 TATTTGATGCTGTTGTCGCAAGCACAGTGAAGAGACATGACCAAGCAATTCATGATCA 1597
DB 156316 TCAATTAATTTCTGATATCATGATTAACAGCAAAAGAAATTTTAATATGCTGCGCACTGGC 156257
QY 1598 GAATGAATAATATCAATTAATGAACCTTGCACTGTTGAACCTTGCGACAGCCTTA 1654
DB 156256 TTAAAAAATTTTAAAGAAAAATGAAACATGCAAAATTAATGATTTAGAGTCCAA 156200

RESULT 15
US-09-679-409-1/c
; Sequence 1, Application US/09679409
; Patent No. 6555316
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel

```

APPLICANT: Blumenfeld, Marca
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
APPLICANT: Essioux, Laurent
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENE, PROTEINS AND BIALLELIC MARKERS
FILE REFERENCE: 53,0515.CIP
CURRENT APPLICATION NUMBER: US/09/679,409
CURRENT FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 09/539,333
PRIORITY FILING DATE: 2000-03-03
PRIORITY APPLICATION NUMBER: 09/416,384
PRIORITY FILING DATE: 1999-10-12
PRIORITY APPLICATION NUMBER: 60/168,088
PRIORITY FILING DATE: 1999-11-30
NUMBER OF SEQ ID NOS: 134
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 319608
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 199122..201122
OTHER INFORMATION: 5' regulatory region
NAME/KEY: exon
LOCATION: 201123..201234
OTHER INFORMATION: exon S
NAME/KEY: exon
LOCATION: 201123..201560
OTHER INFORMATION: exon S2
NAME/KEY: exon
LOCATION: 214676..214793
OTHER INFORMATION: exon T
NAME/KEY: exon
LOCATION: 215702..215746
OTHER INFORMATION: exon U
NAME/KEY: exon
LOCATION: 216836..216994
OTHER INFORMATION: exon V
NAME/KEY: exon
LOCATION: 216836..217077
OTHER INFORMATION: exon V2
NAME/KEY: exon
LOCATION: 217671..217764
OTHER INFORMATION: exon V1
NAME/KEY: exon
LOCATION: 227655..227736
OTHER INFORMATION: exon V4
NAME/KEY: exon
LOCATION: 238715..238919
OTHER INFORMATION: exon V3
NAME/KEY: exon
LOCATION: 240440..240673
OTHER INFORMATION: exon W
NAME/KEY: exon
LOCATION: 240440..241153
OTHER INFORMATION: exon W2
NAME/KEY: exon
LOCATION: 241072..241291
OTHER INFORMATION: exon X
NAME/KEY: exon
LOCATION: 244353..244561
OTHER INFORMATION: exon Y
NAME/KEY: exon
LOCATION: 246273..247802
OTHER INFORMATION: exon Z
NAME/KEY: misc.feature
LOCATION: 247803..249803
OTHER INFORMATION: 3' regulatory region
NAME/KEY: allele
LOCATION: 8316
OTHER INFORMATION: 99-27943-150 : polymorphic base G or C
NAME/KEY: allele

```

1	LOCATION: 21672	OTHER INFORMATION: 99-27935-193	: polymorphic base G or C
2	NAME/KEY: allele		
3	LOCATION: 65485	OTHER INFORMATION: 8-128-33	: polymorphic base C or T
4	NAME/KEY: allele		
5	LOCATION: 95396	OTHER INFORMATION: 99-31960-363	: polymorphic base A or G
6	NAME/KEY: allele		
7	LOCATION: 107281	OTHER INFORMATION: 99-24656-260	: polymorphic base A or G
8	NAME/KEY: allele		
9	LOCATION: 160640	OTHER INFORMATION: 99-24639-163	: polymorphic base A or C
10	NAME/KEY: allele		
11	LOCATION: 160876	OTHER INFORMATION: 99-24634-108	: polymorphic base A or T
12	NAME/KEY: allele		
13	LOCATION: 168974	OTHER INFORMATION: 99-7652-162	: polymorphic base A or G
14	NAME/KEY: allele		
15	LOCATION: 170810	OTHER INFORMATION: 99-16100-147	: polymorphic base A or G
16	NAME/KEY: allele		
17	LOCATION: 173358	OTHER INFORMATION: 99-5862-167	: polymorphic base A or G
18	NAME/KEY: allele		
19	LOCATION: 189957	OTHER INFORMATION: 99-5919-215	: polymorphic base A or G
20	NAME/KEY: allele		
21	LOCATION: 197163	OTHER INFORMATION: 99-24658-410	: polymorphic base A or G
22	NAME/KEY: allele		
23	LOCATION: 200778	OTHER INFORMATION: 8-303-235	: polymorphic base A or G
24	NAME/KEY: allele		
25	LOCATION: 202651	OTHER INFORMATION: 8-300-221	: polymorphic base A or G
26	NAME/KEY: allele		
27	LOCATION: 202679	OTHER INFORMATION: 8-300-193	: polymorphic base A or G
28	NAME/KEY: allele		
29	LOCATION: 203378	OTHER INFORMATION: 8-299-128	: polymorphic base A or T
30	NAME/KEY: allele		
31	LOCATION: 204138	OTHER INFORMATION: 8-296-213	: polymorphic base A or T
32	NAME/KEY: allele		
33	LOCATION: 204605	OTHER INFORMATION: 8-252-190	: polymorphic base C or T
34	NAME/KEY: allele		
35	LOCATION: 204934	OTHER INFORMATION: 99-24644-194	: polymorphic base A or G
36	NAME/KEY: allele		
37	LOCATION: 205206	OTHER INFORMATION: 8-295-248	: polymorphic base A or C
38	NAME/KEY: allele		
39	LOCATION: 205329	OTHER INFORMATION: 8-295-125	: polymorphic base C or T
40	NAME/KEY: allele		
41	LOCATION: 206064	OTHER INFORMATION: 8-293-130	: polymorphic base A or G
42	NAME/KEY: allele		
43	LOCATION: 206545	OTHER INFORMATION: 8-292-198	: polymorphic base A or G
44	NAME/KEY: allele		
45	LOCATION: 207313	OTHER INFORMATION: 8-251-322	: polymorphic base A or G
46	NAME/KEY: allele		
47	LOCATION: 208285	OTHER INFORMATION: 8-289-322	: polymorphic base A or G
48	NAME/KEY: allele		
49	LOCATION: 208960		

Job time : 352 secs

```
OTHER INFORMATION: 8-287-249 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 209123
OTHER INFORMATION: 8-287-86 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 209631
OTHER INFORMATION: 8-285-319 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210361
OTHER INFORMATION: 8-283-278 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210463
OTHER INFORMATION: 8-283-176 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210486
OTHER INFORMATION: 8-283-153 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210583
OTHER INFORMATION: 8-283-56 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 210879
OTHER INFORMATION: 8-282-345 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210964
OTHER INFORMATION: 8-282-260 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 210979
OTHER INFORMATION: 8-282-245 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 211050
OTHER INFORMATION: 8-282-174 : variable motif AAAGG or GAAGGAAGGAAGGAAGA
NAME/KEY: allele
LOCATION: 211132
OTHER INFORMATION: 8-282-92 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 211247
OTHER INFORMATION: 8-281-367 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211315
OTHER INFORMATION: 8-281-299 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211366
OTHER INFORMATION: 8-281-248 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 212520
OTHER INFORMATION: 8-279-197 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 212821
OTHER INFORMATION: 8-278-289 : polymorphic base C or T
NAME/KEY: allele
```

Query Match 2.1%; Score 37; DB 3; Length 319608;

Best Local Similarity 47.3%; Pred. No. 37;

Matches 112; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```
OY 1418 ACAATCAAGGTTCCCATTTAAACTCTCCATGTTACTTCTTACCAACATTT 1477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156436 ACATCTGATGTTCTCACAACAAATTTCCACATTCCTACTGCTGATGAAGACA 156377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1478 GTGTTTCTTATCTAGCCAAAGTATCTATTGAAGTGAACCTTGCCCTCCAAATTAGAATG 1537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156376 GAATTCCTACTATTAACAATGATGATTCACCTGATGCTCAATGATTTTCCACTTTTGG 156317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1538 TATTGATGCTGTGTGCGAAGGCAAGTGAAGAGAACATGACAAAGACATCTAGTCA 1597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156316 TCAATTAATTCGTGTATCATGATTAACAGCAAGAAATTAATATCTGCTGGCACTGGC 156257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1598 GAAATGAAATATCAATTAATGAACCTGCACTGTGAACCTGGCGAGACCTTA 1654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156256 TTAATAAATTTTAGAAAAAATGAACATGCAATTAATTAATTAATAGTCA 156200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 10, 2006, 16:37:10

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 10, 2006, 14:09:32 ; Search time 312 Seconds
(without alignments)
5217.515 Million cell updates/sec

Title: US-10-724-806-4

Perfect score: 2993
Sequence: 1 MPFHEGVAILFLYLFL.....EALPDVDSPEGGSTEDNLQ 580

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1403666 segs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_plus_p2n.model -DEV=xlh
-O=/abs/ABSSWEB_spool/US10724806/runat.10072006.140252.7805/app_query.fasta.1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss07 -USER=US10724806.@CGN_1.1.204.@runat.10072006.140252.7805 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /EMC_Celerra_SIDS3/ptodata/2/ina/1-COMB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5-COMB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A-COMB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B-COMB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7-COMB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H-COMB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCIDS-COMB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP-COMB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE-COMB.seq.*
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2795	93.4	1743	3	US-09-657-252-1 Sequence 1, Appl1
2	1099.5	36.7	1461	5	US-09-974-300-501 Sequence 501, App
3	351	11.7	324	3	US-09-270-767-1720 Sequence 1720, Ap
4	351	11.7	324	3	US-09-270-767-17002 Sequence 17002, Ap
5	318	10.6	2475	3	US-09-799-451-932 Sequence 932, App
6	310	10.4	2028	3	US-10-162-012-28 Sequence 28, Appl
7	310	10.4	2326	3	US-10-162-012-26 Sequence 26, Appl
8	307	10.3	2107	3	US-09-799-451-234 Sequence 234, App

9	298	10.0	1566	3	US-09-540-236-273 Sequence 273, App
10	298	10.0	6703	3	US-09-596-002-7 Sequence 7, Appl1
11	296	9.9	2450	3	US-09-949-016-5012 Sequence 5012, Ap
12	296	9.9	3342	3	US-09-252-991A-7814 Sequence 7814, Ap
13	293.5	9.8	1455	3	US-09-252-991A-7678 Sequence 7678, Ap
14	293.5	9.8	1485	3	US-09-252-991A-7528 Sequence 7528, Ap
15	284	9.5	2238	2	US-07-841-651-1 Sequence 1, Appl1
16	276.5	9.2	1647	3	US-09-543-681A-822 Sequence 822, App
17	274.5	9.2	2394	3	US-09-799-451-931 Sequence 931, App
18	271.5	9.1	2157	3	US-09-657-960-2 Sequence 2, Appl1
19	271.5	9.1	10302	3	US-09-657-960-1 Sequence 1, Appl1
20	271.5	9.1	25356	3	US-09-976-594-750 Sequence 750, App
21	265	8.9	1830121	3	US-09-557-884-1 Sequence 1, Appl1
22	265	8.9	1830121	3	US-09-643-990A-1 Sequence 1, Appl1
23	265	8.9	1830121	3	US-10-158-865-1 Sequence 1, Appl1
24	260.5	8.7	1695	3	US-09-602-789A-495 Sequence 495, App
25	259	8.7	1527	3	US-09-489-039A-370 Sequence 370, App
26	258	8.6	3280	3	US-09-799-451-818 Sequence 818, App
27	254	8.5	1506	3	US-09-328-352-2245 Sequence 2245, Ap
28	245.5	8.2	1557	3	US-09-134-001C-1907 Sequence 1907, Ap
29	244.5	8.2	987	5	US-09-974-300-500 Sequence 500, App
30	243	8.1	1776	3	US-09-540-236-570 Sequence 570, App
31	241.5	8.1	1332	3	US-09-252-991A-10993 Sequence 10993, A
32	241.5	8.1	1656	3	US-09-252-991A-11258 Sequence 11258, A
33	241.5	8.1	2217	3	US-09-252-991A-11190 Sequence 11190, A
34	241	8.1	89047	3	US-09-596-002-34 Sequence 34, Appl1
35	240.5	8.0	1491	3	US-09-489-039A-1601 Sequence 1601, Ap
36	236.5	7.9	1665	3	US-09-543-681A-2714 Sequence 2714, Ap
37	233.5	7.8	1551	3	US-09-252-991A-12931 Sequence 12931, A
38	233.5	7.8	1659	3	US-09-252-991A-13081 Sequence 13081, A
39	233.5	7.8	2073	3	US-09-252-991A-12858 Sequence 12858, A
40	232	7.8	13542	3	US-08-956-171E-154 Sequence 154, App
41	232	7.8	13542	3	US-08-781-986A-154 Sequence 154, App
42	230.5	7.7	1521	3	US-09-543-681A-4074 Sequence 4074, Ap
43	230	7.7	1698	3	US-09-489-039A-1243 Sequence 1243, Ap
44	226.5	7.6	1461	3	US-09-543-681A-2066 Sequence 2066, Ap
45	224.5	7.5	1830121	3	US-09-557-884-1 Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-657-252-1
Sequence 1, Application US/09657252
Patent No. 6500643
GENERAL INFORMATION:
APPLICANT: Wu, Dong-Hai
APPLICANT: Gu, Yunrong
APPLICANT: Millard, William
APPLICANT: He, Yun-Je
TITLE OF INVENTION: Human High Affinity Choline Transporter cDNA
FILE REFERENCE: MEMB00-639
CURRENT APPLICATION NUMBER: US/09/657, 252
CURRENT FILING DATE: 2000-09-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 1743
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1743)
US-09-657-252-1

Alignment Scores:
Pred. No.: 3,42e-314
Score: 2795.00
Percent Similarity: 96.6%
Best Local Similarity: 92.6%
Query Match: 93.4%
Length: 1743
Matches: 537
Conservative: 23
Mismatch: 20
Indels: 0
Gaps: 0

```

US-10-724-806-4 (1-580) x US-09-657-252-1 (1-1743)
OY 1 MetProphHisValGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 1 ATGGCTTCCATGTCGGAAGGAGCTGATGATCATCGTGTCTTACCTTATTTTGGCG 60
OY 21 ValGlyIleTPAlaIleTPLeuTyrLeuSerGlyAspProGluIleArgSerGlu 40
Db 61 GTTGAATATGCGGCGCTCGGAGAACCAAAACAGTGGCAGCGCAGAGAGCGCGAA 120
OY 41 AlaIleIleValGlyIleArgAspIleGlyLeuLeuValGlyIlePheThrMetThrAla 60
Db 121 GCCATCATAGTTGGTGGCCGAGATATGTTATTTGGTTGGATTTTCCATGACAGT 180
OY 61 ThrTPValGlyIleGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db 181 ACCTGGTCGAGGAGGATATATCATGACAGCTGAGGAGGAGTTATATGACAGTTAT 240
OY 81 GlyLeuAlaIleTPAlaIleAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyIleLeu 100
Db 241 GCCCTAGCTTGGGCTCAGGACCAATGGATATCTCTAGTCTGATTTTGGTGGCCCTG 300
OY 101 PhePheAlaIleProMetArgSerLeuGlyTyrValIleThrMetLeuAspProPheLeuGln 120
Db 301 TTCTTTGCAAAACCTATGCGTTCAAGGGGATATGACCATTTAGACCCGTTTCAAGCA 360
OY 121 IleTyrGlyIleArgMetGlyIleLeuLeuPheIleProAlaLeuMetGlyIleMetPhe 140
Db 361 ATCTATGAAAAACGATGGGCGGAGCTCTGTTATTTCCGACAGTATGGGAGAAATGTTT 420
OY 141 TPAlaIleAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 421 TGGGCTGAGCAATTTCTCTGCTTTGGAGCCACCATCGGATCATCATATGGAT 480
OY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyIle 180
Db 481 ATGCACATTTCTGTATCATCTCTGACATTCATTCGACCTCTGACACTGGTGGAGGG 540
OY 181 LeuTyrSerValAlaIleTyrThrAspValAlaGluLeuPheCysIlePheIleGlyLeuTyr 200
Db 541 CTCTATTCGTGGCTACAGCTGATGTCGTTCAGCTCTTTGCAATTTTGTAGGCGCTGG 600
OY 201 IleSerValProPheAlaLeuSerHisProAlaValIleAspIleGlyPheThrAlaVal 220
Db 601 ATGACCGTCCCTTTGCATGTCATCTGACATCTGACATCGGAGATCGGATTCATCTGCTGG 660
OY 221 HisAlaIleTyrGlnSerProTyrLeuGlyThrIleGlnSerValGluValTyrThrTyr 240
Db 661 CATGCCAAATACCAAAAGCGCTGGCTGGGAACTGTTGATCTCATGGAATCTTACTCTTGG 720
OY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleIleProTyrGlnAlaTyrPheGlnArg 260
Db 721 CTGTAGTATGTTCTGTGTTGATGCTGGGGTGAATCCCATGCAAGCATATTCAGAGG 780
OY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db 781 GTTCTCTCTCTCTCTCTGACGACCATATGCTCAAGTGTCTCTCTCTGACACTTCTGGG 840
OY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300
Db 841 TGCCGTGGATGGCCATCCCAAGCATCTCATTTGGGCGCATTTGGACCTCCACAGACTGG 900
OY 301 AsnGlnThrAlaTyrGlyTyrProAspProIleTyrIleGluGluAlaAspMetIleLeu 320
Db 901 AACACAGACTGATATGGGCTTCCAGATCCCAAGACTACAGAGAGGACAGATGATTTTA 960
OY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCAATGTTCTGCAATATCTGCTGCTGATATTTCTTTTGTGTTGTTGTTGTTGAGTT 1020
OY 341 SerAlaIleValMetSerSerAlaAspSerIleLeuSerAlaSerSerMetPheAla 360
Db 1021 TCTGCTGCTGTTATGTCATGACAGATTTCTTCCATCTTGTCCAGCAAGTTCCATGTTTGA 1080

```

```

OY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspIleGluIleValTyrVal 380
Db 1081 CGGAACATCTACAGCTTCTCTTCCAGCAAAATGCTTGGACAAAGAAATCGTTGGGTT 1140
OY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400
Db 1141 ATGCAATCACAGTGTGTTGTTGGACCATCTGCACAGCCAGTGGCTGTCAGAA 1200
OY 401 ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
Db 1201 ACTGTATGGGCTCTGCTACTCATGTTCTGACCTGTTTACATGTTATCTTCCCCAG 1260
OY 421 LeuLeuCysValLeuPheIleLeuGlyIleThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db 1261 CTGCTTGTGTACTCTTGTTAAGGAAACCAACCTATGGGCGCTGGCAGGTTATGTT 1320
OY 441 PheGlyLeuPheLeuArgIleThrGlyIleGluProTyrLeuTyrLeuGlnProLeuIle 460
Db 1321 TCTGGCTCTTCTGAGATTAACCTGAGAGGAGCCATATCTGTATCTTCAGCCCTGATC 1380
OY 461 PheTyrProGlyTyrTyrSerAspIleGlyIleTyrAsnGlnArgPheProPheIle 480
Db 1381 TTCTACCTGCGTATTTACCTGATGATATATGTTATATATCAAAATTTCCATTTAAA 1440
OY 481 ThrLeuSerMetValIleThrSerPheThrAsnIleCysValSerTyrLeuAlaIleTyr 500
Db 1441 ACACCTGCAAGTGTATCATCTTTTAAACCAACATTCACCTCTCATCTGACCAAGTAT 1500
OY 501 LeuPheGluSerGlyTyrLeuProProIleAspValPheAspAlaValAlaIleArg 520
Db 1501 CTATTTGAAAGGAACTTTCGACCTAAATTAAGTATTTGATGCTGTTGTCAGAA 1560
OY 521 HisSerGluGluAsnMetAspIleThrIleLeuValArgAsnGluAsnIleIleLeuAsn 540
Db 1561 CACAGTGAAGAAACATGATGATGACATTTCTTGTCAAAATTAATTAATTAATGAT 1620
OY 541 GluLeuAlaProValIleProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIle 560
Db 1621 GAACCTTGACCTGTGTAACCAAGCAGACATGACCTCAGCTCAACTTTCACCAATAAA 1680
OY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyTyrGluAspAsnLeuGln 580
Db 1681 GAGGCTTCTTGTATGTTATTCAGTCCAGAGGCTCGGAGCTGGAAGATTAATTAACA 1740

RESULT 2
US-09-974-300-501
; Sequence 501, Application US/09974300
; Patent No. 7018784
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085-500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 501
; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-501

Alignment Scores:
Pred. No.: 2,63e-117 Length: 1461
Score: 1099.50 Matches: 232
Percent Similarity: 62.84 Conservative: 89

```

Best Local Similarity: 45.4% Mismatches: 147
 Query Match: 36.7% Indels: 43
 DB: 5 Gaps: 8

US-10-724-806-4 (1-580) x US-09-974-300-501 (1-1461)

```

QY 4 HisValGlyGlyLeuValAlaIleIleLeuPheTyrIleLeuIlePheLeuValGlyIle 23
DB 2 CATGGGGGGGGCTTTATCTCCATGATGATATTTCTTGGCTTACCATATTAATTAGGCGCT 61
QY 24 TTPAlaAlaIleTyrIleThrIleAsnSerGlyAsnProGluIleArgSerGluAlaIleIle 43
DB 62 TTTTACGCGCAAGAAATCAAAATCTTTTCCGAT-----ATGATC 103
QY 44 ValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAlaThrTPVal 63
DB 104 GTGGCAAAAGAGCATGCCGTTTTCGTGGATGCTGTGCGATGGCCGCAACATGGGCT 163
QY 64 GlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAla 83
DB 164 GCGCGCGGATATTTAAAGGAGCGCGCATGACCTTACAGCAC-----GGCTCATC 217
QY 84 TTPAlaAlaIleProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeuPheAla 103
DB 218 TGGGCCCAAGCGCTTGGGCTACGATGAGCTGATTAACGCGGTATTTCTTCCGCC 277
QY 104 LysProMetArgSerIleGlyTyrValThrMetLeuAspProPheGlnIleTyrGly 123
DB 278 AGAAAAATCGCGCGTCATCAATTATGACCATTAATGATCTCTCGAACAGCGCTTCGGT 337
QY 124 LysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPheTPAlaAla 143
DB 338 AGCGGATGGCGCGCTTTTATATACCGCGCTTTAGAAATGTTTTCGAGCGCC 397
QY 144 AlaIlePheSerIleLeuGlyAlaThrIleSerValIleIleAspValAspValAlaIle 163
DB 398 GCGATCTTAACGCGCATTTGGGCAACGACTTTCGAAATGATTCGAAATTCGAAATTCGAAACG 457
QY 164 SerValIleValSerIleValIleAlaIleLeuTyrThrLeuValGlyGlyLeuTyrSer 183
DB 458 TCGATTATTTCTTGGCGGATGATGCGCATGCGCTTAACGCTGGCGGAGCGCATGTGGCA 517
QY 184 ValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTPileSerVal 203
DB 518 GTTGCTTACAGATCTCTTCAATGATGATGCTTTCGCGGCTGTTTATAGCGTC 577
QY 204 ProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaValHisAlaLys 223
DB 578 CCATTGTGATGTGCGAATGTCGTGCTCGAT-----ACGATATGGCAAAAT 625
QY 224 Tyr-----GlnSer 226
DB 626 TACAGGCGCATTTCCGACAGCGCCCAATTCGCTTCGCGCTTGAACGGCTGGAAAAAC 685
QY 227 ProTyrLeuGlyThrIleGlnSerValGluValTyrThrTyrLeuAspAsnPheLeu 246
DB 686 CCGGATTTGGGAAACTG-----TTTGGAACTGGTGGGATATATGGCTTCCTC 733
QY 247 LeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArgValLeuSerSerSer 266
DB 734 TTAATTTTCGAGAGTTCATCGATGCGAGTGTACTTTCAGGCGCTTTCGCGCAAAATCG 793
QY 267 AlaThrTyrAlaGlnAlaLeuSerPheLeuAlaIlePheGlyCysLeuValMetAlaLeu 286
DB 794 GAAAGCGCGCATTTGGCGAGTCGATATATGCGGATGATCTGATTCGCGCAT 853
QY 287 ProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyrAsnGlnThrAlaTyrGly 306
DB 854 CCGTGGTATATCATCGAGCTGCGGAAACAGTACCGATTGGAGCGCTTTCGAGGAGAG 913
QY 307 TyrProAspProLysThrIleArgGluGluAlaAspMetIleLeuProIleValLeuGlnTyr 326
DB 914 GCTCCGAGT-----AACCGCGCATATTTTCCGCAAAACGCTTGGCTAT 958

```

```

QY 327 LeuCysProValTyrIleSerPhePheGlyLeuGlyAlaValSerAlaValMetSer 346
DB 959 TTAGCGCCAGGAATCATTCGACAGCCCTTCGCTGGGTCATTCGACAGCCGCTCAT 1018
QY 347 SerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIleTyrGlnLeu 366
DB 1019 AGCATGACTCATCATCTTATTCGATCATCATGAGCGCGCATGAAATTTTACCTCGC 1078
QY 367 SerPheArgGlnAsnAlaSerAspLysGluIleValTyrValMetArgIleThrValLeu 386
DB 1079 CTCATCAACCGCAAGGCGCAACCAAAACGCTGCAAAAAGTGTCAACGCTCAATCAT 1138
QY 387 ValPheGlyAlaSerAlaThrAlaMetAlaLeuMetThrLysThrValTyrGlyLeuTP 406
DB 1139 TTGTTGCGCGCGGAGACCGGTCACTCGGCTGAATGCAAAAGCTTTTACTTATG 1198
QY 407 TyrLeuSerSerAspLeuValTyrIleIleIlePheProGlnLeuLeuCysValLeuPhe 426
DB 1199 TATTTGCTTCGATTTATTTATTCATTTCTTTTCCAGTTAACATGGCCCTCTT 1258
QY 427 IleGlyGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArg 446
DB 1259 TATTAAGCAAAATCTTTACCGGTGATTCGATTCGATTCGATTCGATTCGATTCGAG 1318
QY 447 IleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIlePheTyrProGlyTyrTyr 466
DB 1319 CTCGGCGGTGTGAACCCGATTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTC 1375
QY 467 SerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLysThrLeuSerMetValThr 486
DB 1376 GAAGACGACGAGCTCTC-----TTTCTTTCGAGACGCTGGTGACGATCACA 1423
QY 487 SerPhePheThrAsnIleCysValSerTyrLeu 497
DB 1424 GCATTTATGACGATCTTCGCTGTGCGAATTG 1456

```

RESULT 3
 US-09-270-767-1720
 ; Sequence 1720, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1720
 ; LENGTH: 324
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 ; US-09-270-767-1720

Alignment Scores:
 Pred. No.: 2 05e-31 Length: 324
 Score: 351.00 Matches: 69
 Percent Similarity: 81.3% Conservative: 18
 Best Local Similarity: 64.5% Mismatches: 20
 Query Match: 11.7% Indels: 0
 DB: 3 Gaps: 0

US-10-724-806-4 (1-580) x US-09-270-767-1720 (1-324)

```

QY 266 SerAlaThrTyrAlaGlnAlaLeuSerPheLeuAlaIlePheGlyCysLeuValMetAla 285
DB 3 ACGGCAAGAGGCGCCAGCTTCTGCTATGTTGACGCGCGGAGCATTTTGAAGGCC 62
QY 286 LeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyrAsnGlnThrAlaTyr 305
DB 63 ATTCCCGGTGCTCATTCGAGGATTCGCAAGGCTTACACCTTGGAACAGACGATTCAC 122

```


Oy	306	GIlyTyrProAspProLysThirHygIuAlaAspMetIleuProlIeValLeuGln	325
Db	123	AAGGACCCCTATCCCCCTGACCCGTGACGACGACGACGATGATTCTGCCCATGGTGTCCGAC	182
Oy	326	TyrIleuCySProValTyrIleSerPhePheGlyLeuGlyAlaValaIserAlaAlaValMet	345
Db	183	TACCCACAGCCCTGACCTCGTGTCTCTTGTGGATTGGGGCGCTGTTCGCCGCCGGATG	242
Oy	346	SerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIleTyrGln	365
Db	243	TCCTCCGCCGACTCTCCGTGCTCTTCGCCCGCTCCATGTTCCCTCGAAGCTGTACAA	302
Oy	366	LeuSerPheArgGlnAsnAla	372
Db	303	TTGATTTCCTCGTCAGAAAGCGC	323
RESULT 4			
US-09-270-767-17002			
; Sequence 17002, Application US/09270767			
; Patent No. 6703491			
; GENERAL INFORMATION:			
; APPLICANT: Homburger et al.			
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster			
; FILE REFERENCE: File Reference: 7326-094			
; CURRENT APPLICATION NUMBER: US/09/270,767			
; NUMBER OF SEQ ID NOS: 62517			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 17002			
; LENGTH: 324			
; TYPE: DNA			
; ORGANISM: Drosophila melanogaster			
US-09-270-767-17002			
Alignment Scores:			
Pred. No.:	2.05e-31	Length:	324
Score:	351.00	Matches:	69
Percent Similarity:	81.3%	Conservative:	18
Best Local Similarity:	64.5%	Mismatches:	20
Query Match:	11.7%	Indels:	0
DB:	3	Gaps:	0
US-10-724-806-4 (1-580) x US-09-270-767-17002 (1-324)			
Oy	266	SerAlaThrTyrAlaGlnValLeuSerPheLeuAlaAlaPheGlyCysIeuValMetAla	285
Db	3	ACGGCAGGAAGGCGCCAGCTTGTGTCTATGTTGACGCCCGGAGTCATTTGATGGCC	62
Oy	286	LeuProlAlaIleCysIleGlyAlaIleGlyAlaSerThrAspIhrpAsnGlnThrAlaTyr	305
Db	63	ATTCGCCCGTGTCTATCGGAGCGATTGCCAAGGCTACACCTTGAGACGACGACGATTAC	122
Oy	306	GlyTyrProAspProLysThirHygIuAlaAspMetIleuProlIeValLeuGln	325
Db	123	AAGGACCCCTATCCCCCTGACCCGTGACGACGACGACGATGATTCTGCCCATGGTGTCCGAC	182
Oy	326	TyrIleuCySProValTyrIleSerPhePheGlyLeuGlyAlaValaIserAlaAlaValMet	345
Db	183	TACCCACAGCCCTGACCTCGTGTCTCTTGTGGATTGGGGCGCTGTTCGCCGCCGGATG	242
Oy	346	SerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIleTyrGln	365
Db	243	TCCTCCGCCGACTCTCCGTGCTCTTCGCCCGCTCCATGTTCCCTCGAAGCTGTACAA	302
Oy	366	LeuSerPheArgGlnAsnAla	372
Db	303	TTGATTTCCTCGTCAGAAAGCGC	323
RESULT 5			
US-09-799-451-932			
; Sequence 932, Application US/09799451			
; Patent No. 6783969			
; GENERAL INFORMATION:			

	APPLICANT:	Tang, Y. Tom	
	APPLICANT:	Zhou, Ping	
	APPLICANT:	Goodrich, Ryle	
	APPLICANT:	Aseundi, Vinod	
	APPLICANT:	Ren, Feiyun	
	APPLICANT:	Zhang, Jie	
	APPLICANT:	Xue, Aidong J.	
	APPLICANT:	Zhao, Qing A.	
	APPLICANT:	Wang, Jian-Rui	
	APPLICANT:	Ma, Yunting	
	APPLICANT:	Yamazaki, Victoria	
	APPLICANT:	Chen, Rui-hong	
	APPLICANT:	Wang, Zhiwei	
	APPLICANT:	Yang, Yonghong	
	APPLICANT:	Wehrman, Tom	
	APPLICANT:	Ghosh, Reena	
	APPLICANT:	Drimac, Radcoje T.	
	TITLE OF INVENTION:	No. 6783969e1 Nucleic Acids and Peptideptides	
	FILE REFERENCE:	803	
	CURRENT APPLICATION NUMBER:	US/09/799,451	
	CURRENT FILING DATE:	2001-03-05	
	NUMBER OF SEQ ID NOS:	948	
	SOFTWARE:	pc_FL_genes Version 2.0	
	SEQ ID NO 932		
	LENGTH:	2475	
	TYPE:	DNA	
	ORGANISM:	Homo sapiens	
	FEATURES:		
	NAME/KEY:	CDS	
	LOCATION:	(347)..(2371)	
	US-09-799-451-932		
	Alignment Scores:		
	Pred. No.:	4,04e-26	Length: 2475
	Score:	318.00	Matches: 148
	Percent Similarity:	39.0%	Conservative: 106
	Best Local Similarity:	22.7%	Mismatches: 224
	Query Match:	10.6%	Indels: 174
	DB:	3	Gaps: 27
QY	US-10-724-806-4 (1-580) x US-09-799-451-932 (1-2475)		
Db	8 LeuValAlaIlelleuPheTyrlenuilePheUValGlylletPrAlaIatrp	27	
Qy	ATGGCGGTGTAGTTCTTGAATTCTTCCTTGTCGGCTGTGGACATAGGTGCACAGTg	490	
Db	431 ATGGCGGTGTAGTTCTTGAATTCTTCCTTGTCGGCTGTGGACATAGGTGCACAGTg	490	
Qy	28 lyeThrlysAsnserclYasnProglugluArgergluAlalleleValGlylYArg	47	
Db	491 AAAGCACA-----AGAGCACAGTAAGGACTTCTGCGTAGGAGG	535	
Qy	48 AspIleGlYleuUValGlylYpHeThmetThralarHrPrVaIGlylYtyr	67	
Db	536 GACATGGTGGGAGCCAGTGGTGCATCCTGGTGGCACAAATGTTGGAAAGTGACAT	595	
Qy	68 IleasnglyThralaeluAlvalYtelyProglyCysglyleuAlatPrAlahisAla	87	
Db	596 TTCATTGGCGCTG-----CGAGGTGAGGTGCTGCTACGGCATTTCT	637	
Qy	88 ProIleGlYTYserleuser-----leuIleauglylYleuPhe	102	
Db	638 GTATCAGCTTAAATTAATAGCTGTGTTTCTGTCGATGATGCTGGATCTTCTA	697	
Qy	103 AlalysProMetArgerSerlysglyTYrValThrMetleuaapPropheylsglnIetyr	122	
Db	698 -----CCCATCACTTCTGTCAGGTGCACAAAGATCCAGAATCTTACCG-	745	
Qy	123 GlylsArMeGlYglyleuUheHlePrAlaleuMetGlylumePhe-----	140	
Db	746 ---AAAGGCTTGGTGGCATCAGA---ATCCCATCATCTTGGCTGACTTACTTATT	799	
Qy	141 -----TPrlalaAlaIlePheSerAlau	149	


```

      / PRIOR FILING DATE: 2001-03-28
      / PRIOR APPLICATION NUMBER: US 10/109,029
      / PRIOR FILING DATE: 2002-03-28
      / PRIOR APPLICATION NUMBER: PCT/US02/09728
      / PRIOR FILING DATE: 2002-03-28
      / PRIOR APPLICATION NUMBER: US 60/290,288
      / PRIOR FILING DATE: 2001-05-11
      / PRIOR APPLICATION NUMBER: US (not assigned)
      / PRIOR FILING DATE: 2002-05-13
      / NUMBER OF SEQ ID NOS: 48
      / SOFTWARE: FastSeq for Windows Version 4.0
      / SEQ ID NO 28
      / TYPE: DNA
      / LENGTH: 2028
      / ORGANISM: Homo sapiens
US-10-162-012-28

Alignment Scores:
Pred. No.:      2,47e-25      Length:      2028
Score:          310.00      Matches:      149
Percent Similarity: 40.2%      Conservative: 111
Best Local Similarity: 23.0%      Mismatches:  224
Query Match:      10.4%      Indels:      164
DB:                3      Gaps:          27

US-10-724-806-4 (1-580) x US-10-162-012-28 (1-2028)

QY      8 LeuValAlaIleIleLeuPheTyrlLeuLeuIlePheLeuValGlyIleTPrAlaIaItrp 27
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      85 ATCGCGGCTGATCTGATCTTGAATTCCTCTTGTCTCGCTGCTGTGACTATGTCACAGTG 144
QY      28 LysThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyIArg 47
      |||||||
DB      145 AAGACCAAA-----ACAGACAC-AGTGAAGGCTACTTCTGGCTGAAG 188
QY      48 AspIleGlyLeuLeuValGly-GlyPheThrMetThrAlaThrTryValGlyIleGlyTy 67
      ::::::::::::::::::::
DB      189 GAACATGGTGGTGGCCAGGTGGTGCATCTTGTGTTGCCAGAAATTTGAAGTGGACA 248
QY      67 rIleAsnGlyThrAlaGluAlaValTyrgIyProGlyCyGlyLeuAlaTPrAlaHisAl 87
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      249 TTTCATGGCCCTGGCA-----GGTCAAGTGCTGCTACGGGCATTC 290
QY      87 abroIleGlyTyrrSerIeuser-----LeuIleLeuGlyIleuPhePh 102
      : : : : :
DB      291 TGTATCAGCTTATGAATTAATGCTGTTTCTGTGCTGATGTTGGCTGATCTTCTC 350
QY      102 eAlaLysProMetArgSerIleGlyTyrrAlaThrMetLeuAspProPheLysGlnIeTy 122
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      351 A-----CCATTTACATGCTGTGTACAGGCACACAGATCCAGAAATCCTACGG--- 399
QY      122 rGlyLysArgMetGlyIleuLeuPheIleProAlaLeuMetGlyIleuPheTPrAl 142
      : : : : :
DB      400 ---AAGCGCTTGGTGGCATCAGA--ATCCCATCATCTCGGTGTACTTACCTATT 452
QY      142 aAlaIaIlePheSerAlaLeuGlyAlaThrIle-----SerValIleIleAsp- 158
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      453 TATCTACATCTTCACCAAGATCTCCGGTAGACATGTATGACGGTGCCATCTTCATCCAGCA 512
QY      159 -----ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTy 175
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      513 GCTTTCGCACCTGGATCTGTACTCTGGCCATAGTGGGTACTGGCCATCACTGCTGTATA 572
QY      175 rThrLeuValGlyIleuTyrrSerValAlaTyrrThrAspValAlaGlnLeuPheCysIl 195
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      573 CAGGGTGTGGTGGGCGCTGCTGTGATCTACACGGAGTCCCTCAGACCGCTGATCAT 632
QY      195 ePheIleGlyLeuTPrIleSerValProPheAlaLeuSerHisProAlaAlaThrAspIl 215
      : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      633 GCTTATAGAGCGGCTCACTTGATGGGCTAC-----AGTTTGGCGGGGTGGT--- 681
QY      215 eGlyPheThrAlaValHisAlaLysTyrr----- 224

```

Db	682	GGGATGAGAGACTGAAAGGAAGTACTCTTGGCCCTGGCTAGCAACCGAGACTGAGA	740
Qy	224	-----	224
Db	741	CAGCAGCTGCGGGGTGCCCCGGGAGAATGCTTCATATTTCGAGATCGCTGATC	800
Qy	225	-GlnSerProTrpLeuGlyThrIleGluSerValGluVal-----TyrThrTrpLe	241
Db	801	TGATCTCCCGGGGGGGGGGCTCTATTGGAAATGTCATCCCATCCCTCGTATCTGTG	860
Qy	241	uASPasnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrPheGlnArgVa	261
Db	861	CACGGAT-----CAGGTGATTGGCCGCGGAC	887
Qy	261	lleuSerSerSerSerIleThrTyrAlaGln-----ValLeuSerPheLeuAl	277
Db	888	TCGTGGTCCCAAGAACCTGTCCTCAATGCCAAGAGAGGTGCTCTGATGGCTGATATCCGAA	947
Qy	277	alaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSe	297
Db	948	GGTGTGCGCCCTCTTCATTAATGGTTCCT-----GGANNGGTCACCGCAT	995
Qy	297	rThraspTrpAsnGlnThrAlaTyrGlyTyrProASPProLysThrIysGlnGlu-----	315
Db	996	CCTCTTCCCATGATCAGTGCCTGTGCA-----GATCCAGAGATCTGCCAAGAAATCTG	1049
Qy	316	-----AlaAspMetIleLeuProIleValLeuGlnTyrLeuCysPr	329
Db	1050	CAGCAACCCCTCAGGCTGTTCGGACATCGGCTATCCAACTGCTGTCGAACTCTGCCC	1109
Qy	329	oValTyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaValMetSerSerAlaAs	349
Db	1110	CACAGGCTCGTGGGTGATGATGAGTGGTGGTGGTGGCTCTCATGTCCTCCCTCAC	1169
Qy	349	pSerSerIleLeuSerIleAspSerMetPheAlaArgAsnIleTyrGlnLeuSerPheAr	369
Db	1170	CTCATCTTTAACTGTCAGACACATCTTACCATGGAACCTCTGGAAT--CACCTCCG	1226
Qy	369	gGlnAsnAlaSerAspLysGlnIleValIleTrpValMetArgIleThrValLeuValPheG	389
Db	1227	GCCTCGGGCATCTGAGAAAGAGCATGATGTGGGACGGGTGTTTGG-----	1275
Qy	389	YAlaSerAlaThrAlaMetAlaLeuLeuThrIysThrValTyrGlyLeuTrp-----	406
Db	1276	-----CTGCTGTGGTCTGGTCTCCATCTCTGGATCCCTGT	1313
Qy	407	-----TyrLeuSerSerAspLeuValTyrIl	415
Db	1314	GGTCCAGGCCACCGAGGGCGGCAGCTTTCATCTATATACAGTCCATCAGCTCCTTACT	1373
Qy	415	eIleIlePheProGlnLeuLeuCysValLeuPheIleIysGly-----Th	430
Db	1374	G-----CAGCGCTGTGGGGGTGTCCTTCATCATGGAGATTTTCTGGAAGAGAC	1424
Qy	430	rAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArgIleThrGly--	449
Db	1425	CAATGAAAGAGGTGCTTTCGGGGCTGATGTCCGGGCTGCTCTGGGCTTGGTAGGCT	1484
Qy	450	-GlyLeuProTyrLeuTyrIleGlnProLeuIlePheTyrProGlyTyrTyrIserAspLy	469
Db	1485	GGTCTGTGACCTTATTATTTCGTGACGCTCGATGGAGCACAGCA-----GATGA	1532
Qy	469	sAsnGlyIleTyrAsnGlnArgPheProPheLeuYThrLeuSerMetValThrSerPhePh	489
Db	1533	GCGCCGGGTCTCGGTGAAGACATCTACTTACTTCCATGATCCTGTCCACGGT	1592
Qy	489	eThrAsnIleCysAlaSerTyrIleAlaAlaTyrLeuPheGluSerGlyThrLeuProPr	509
Db	1593	CACCTTATCACTGTCTCCACCGGTAGGTGTT-----ACAGAGCACAC	1637
Qy	509	oLysLeuAspValPheAspAlaValAlaAlaArgHisSerGlnIleuAsnMetAspLysTh	529
Db	1638	CTCCACAGAGATGGTCAAGCCACCTGG-----AC	1664


```
OY 261 lLeuSerSerSerAlaThrTyrAlaGln-----ValLeuSerPheLeuAl 277
|||:||||: ||: ||||:||||: |||: ||||:||||:
Db 1065 TCTGCTGTCAGAAAGCCTGTCATGAGCAAGAGAGAGTCTGATGAGCTGATCTACTGAA 1124
OY 277 aAlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSe 297
1125 GATGCTGCCCCCTTTCATTAATGATGTTCCCT-----GGATGATGAGCCGCAT 1172
OY 297 rThraSPTrpaSngInThraAlaTyrGlyTyrProAspProIlySthlySgluGlu----- 315
|||:||||: |||: ||||:||||: |||: ||||:||||:
Db 1173 CCTCTTCCAGATCAAGTGGCTGTGCA-----GATCCAGAGATCTGCCAAGAATCTG 1226
OY 316 -----AlaAspMetIleuProIleValIleuGlnTyrLeuCysPr 329
1227 CAGCAACCCCTCAGAGCTGTGGACATGCGCATGCTCCAAACCTGCGGAACTCCCTG 1286
OY 329 oValTyrIleSerPhePheGlyLeuGlyAlaValSerAlaValMetSerAlaAs 349
1287 CACAGAGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1346
OY 349 pSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIleTyrGlnLeuSerPheAr 369
1347 CTCATCTTTTAAAGTCCAGACCATCTTCAACATGACCTCTGGAAT---CACCTCG 1403
OY 369 gGlnAsnAlaSerAspIlySgluIleValTyrValMetArgIleThrValLeuValPheG 389
1404 GCCTGGGCATCTGAGAGAGACCTCATATTTGGGACAGGAGTTTGTG----- 1452
OY 389 yAlaSerAlaThrAlaMetAlaLeuThrIlySthValTyrGlyLeuTyrP----- 406
1453 -----CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1490
OY 407 -----TyrLeuSerSerAspLeuValTyrI 415
Db 1491 GATCCAGGACAGCCAGGCGGACGCTCTTCACTATATACAGTCCATCAGCTCCACT 1550
OY 415 eIleIlePheProGlnIleuLeuCysValIleuPheIleuSgly-----Th 430
1551 G-----CAGCCGCTGTCGCGGTGCTTTCATCATGATGATGATGATGATGATGATGATG 1601
OY 430 rAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArgIleThrGly-- 449
1602 CAATGAAAAGGATGCTTCTGAGGCTGATCTGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1661
OY 450 -GlyGlnProTyrLeuTyrLeuGlnProLeuIlePheTyrProGlyTyrTyrSerAspI 469
1662 GGTCTGAGCTTATTTAGTGCAGCTGATGATGATGATGATGATGATGATGATGATGATG 1709
OY 469 sAsnGlyIleTyrAsnGlnArgPheProPheIlySthIleuSerMetValThrSerPhe 489
Db 1710 GCGCCCGGCTCTGAGAGAGCATTCACCTTACTTCTCCATGATCTGCTCCACGGT 1769
OY 489 eThraSnlleCysValSerTyrLeuAlaIlySthIleuPheGlnSerGlyThrLeuPro 509
1770 CACCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1814
OY 509 oLysLeuAspValPheAspAlaValAlaArgHisSerGlnGluAsnMetAspIlySth 529
1815 CTCCAAAGAGATGTCAGCCACTG-----AC 1841
OY 529 rIleLeuValArgAsnGluAsnIleLysLeuAsnGluLeuAlaProValIlyProArg 549
1842 CTGTTTACTGCTCAGACCCGCTGCTCCAGAGAAACAGACACACAGCAGCTCCCTT 1901
OY 549 nSerLeuThrLeuSerSerThrPheThraSnllySgluAlaLeuAspValAspSer 569
1902 GTCTCTTACCTCTCAG-----AACGGGAT 1928
OY 569 rProGlnGlySerGlyThr 575
Db 1929 GCCAGAGGCCAGCAGCAGC 1947
+RESULT 8
```

```
US-09-799-451-234
; Sequence 234, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Yang, Yonghong
; APPLICANT: Weinman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799, 451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pl_FL_genes Version 2.0
; SEQ ID NO 234
; LENGTH: 2107
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2057)
; US-09-799-451-234

Alignment Scores:
Pred. No.: 5,87e-25 Length: 2107
Score: 307.00 Matches: 130
Percent Similarity: 39.2% Conservative: 85
Best Local Similarity: 23.7% Mismatches: 192
Query Match: 10.3% Indels: 142
DB: 3 Gaps: 22

US-10-724-806-4 (1-580) x US-09-799-451-234 (1-2107)
OY 11 lIleIlePhePheTyrLeuIlePheLeuValGlyIleTyrAlaAlaTyrIlySthIly 30
|||:||||: |||: ||||:||||: |||: ||||:||||:
Db 174 ATGTCATCTATTTCTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 230
OY 31 AsnSerGlyAsnProGlnGluArgSerGluAlaIleValGlyIlyArgAspIleGly 50
231 AACGAGGATCT-----ATAGAGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 278
OY 231 lLeuLeuValGlyIlyPheThrMetThrAlaThrTyrValGlyIlyGlyTyrIleAsn 70
279 TGGTGGCCGATGGGCGCTCTCTTTGSCAGTAACTGGCAGACACCACTATGTGGGG 338
OY 71 ThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaTyrPheAlaProIleGly 90
339 CTGGGT-----GGACAGAGCGCTTCAAGAGTGCACACCGCTAAC 380
OY 91 TyrSer-----LeuSerIleLeuGlyIlyLeuPhePheAlaIlyPro 105
|||:||||: |||: ||||:||||: |||: ||||:||||:
Db 381 TTGTAATGACTTCTCTAGTATGTGCTGATCTTGGTGGATGATTTGTGCTGCTGCTGCT 440
OY 106 MetArgSerIlySgluTyrValThrMetLeuAspProPheIlySgluIleTyrGlyIly 125
441 ATCAAGTCG---GGGTGATGACCATGCCGGAATATCTCAAG-----AACGGG 485
OY 126 MetGlyGly-----LeuLeuPheIlePro 133
```

```

Db      486 TTTGGGAGAGAGAGCTCCAGGTCTACTCTCCATCTCTCCCTCTTCACTGTGTGGTT 545
Qy      134 AlaIeMeCylGluMeTherPhaIaIaIlePheSerAlaLeuGlyAlaThrIle 153
Db      546 TTGTTAATTTCGACAGACATATTTCGTGAGCCATTTTC-----ATC 587
Qy      154 SerValIleIleAspValaIaIleSerValaIleSerValaIleSerAlaLeuIleAlaIle 173
Db      588 AAGCTGGCTTGGAGATGGACCTTACCTGGCAATCTTACCTCTGGCTAGAGTGGCT 647
Qy      174 LeuTyrThrLeuValGlyLeuTyrSerValaIaIaTyrThrAspValaGluLeuPhe 193
Db      648 GTTACACACCACACTGGGGCTTGGCTGGTATTTACAGACACCCCTCCAGACCATC 707
Qy      194 CysIlePheIleGlyLeuThrIleSerValaIleSerValaIleSerAlaLeuSerIleProAlaValThr 213
Db      708 ATCAGCTGATTTGGCTTTTATCTCATGGGGTTTGCATTATACGAACTTGA----- 761
Qy      214 AspIleGlyPheThrAlaValaIleHisAlaLysTyrGlnSerProThrLeuGlyThrIleGlu 233
Db      762 -----GGTTATAGAGCTTTACCGAAGTACGTAAATGCCACCCATCCGATCGAG 815
Qy      234 -----SerValaGluValTyrThr---TripleAspAsnPheLeu 246
Db      816 GGGGACACTTGACATACATGCTGCTGCTACACACTCGGGCGGACTCTTCCACATC 875
Qy      247 Leu-----MetLeuGlyIleProThr----- 254
Db      876 TTCGAGATGCTGTGACTGGGACATTGCGACGAGAAATTATATTGGATGCCATT 935
Qy      255 -----GlnAlaTyrPheGlnArgValSerSerSer 265
Db      936 ACAGCTTTGGTGTACTGTCGACCAATCAGGTCTATGTGCGAGGCTGCTGTGGCAAG 995
Qy      266 SerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIaPhe----- 279
Db      996 GACATGCTCAGCTGAAGCCGCTGCAATTATGTGCTTACCTGAAGCTGTGCCATG 1055
Qy      280 -----GlyCys 281
Db      1056 TTCCTCATGTGTATGCGGGATGATCAGCCGATCTGTACACAGATATGTAGCATGT 1115
Qy      282 LeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrpAsn 301
Db      1116 GTGGTA-----CCTTCTGAATGCGTAACACACTGTGC-----GTTATGTTGGC 1160
Qy      302 GlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetIleLeuPro 321
Db      1161 TGCACCACTACGCATACCC-----ACGATGTGTGGAA 1196
Qy      322 IleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaValSer 341
Db      1197 CTGATCCCCCAAGACTGCGAGGCTGATGCTTTG-----GTCATGCTG 1241
Qy      342 AlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArg 361
Db      1242 GCTCTCTCATAGCTCCCTGACCTTCATCTTCAACAGCCGACCCCTTCCACATT 1301
Qy      362 AsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrpValMet 381
Db      1302 GACCTCTAC---ACCAAGATGCGGAAGCAAGCTCGGAAGAAAGAGCTCGATAGCTGGA 1358
Qy      382 ArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLysThr 401
Db      1359 CGGATATTTGTTCTCTATTAACCTGTGAGCATGTGTGGTCCCATCGGTACAGATT 1418
Qy      402 ValTyrGly-----LeuTrpTyrLeuSerSerAspLeuValTyrIle-----Ile 416
Db      1419 TCCCAAAATGAGCAACATTCATTCACAGATCAATTTCTAGCTACCTGGGCTCCCA 1478
Qy      417 IlePheProGlnLeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaVal 436

```

```

Db      1479 ATTGACGTCTTTGTGCTTCCATCTTGTAAAAAGATGAATGAAGAGACATTG 1538
Qy      437: AlaGlyTyrIlePheGlyLeuPheLeu----- 445
Db      1539 TGGGGCTATAGTTGTTGACCTTGCATGCGCTTCATTCGATGATTAACAGAGTTGCTTAT 1598
Qy      446 -----ArgIleThrGlyGlyLeuProTyr 453
Db      1599 GGAACAGGAGATTGCTGGCTCCAGTACATGTCACAGATATATCTGTGAGTGCATAT 1658
Qy      454 LeuTyrLeuGlnProLeuIlePheTyr 462
Db      1659 CTGTACTTTCCATGCTCTCTTTT 1685

RESULT 9
US-09-540-236-273
/ Sequence 273, Application US/09540236
/ Patent No. 6673910
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2005-001
/ CURRENT APPLICATION NUMBER: US/09/540,236
/ NUMBER OF SEQ ID NOS: 3840
/ SEQ ID NO 273
/ LENGTH: 1566
/ TYPE: DNA
/ ORGANISM: M. catarrhalis
/ US-09-540-236-273

Alignment Scores:
Pred. No.: 3,996-24 Length: 1566
Score: 298.00 Matches: 132
Percent Similarity: 41.8% Conservative: 99
Best Local Similarity: 23.9% Mismatches: 223
Query Match: 10.0% Indels: 100
DB: 3 Gaps: 18

US-10-724-806-4 (1-580) x US-09-540-236-273 (1-1566)
Qy      9 ValAlaIleIleLeuPheTyrLeuLeuIlePheLeuValGlyIleTrpAlaIaIaTrpLys 28
Db      97 ATTGCTTGGCTGCTCATTTATTTGATGATTCGATTCGAAATTTATGCTTATTTAA 156
Qy      29 ThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyIaIaGasp 48
Db      157 CAAAAAATGAC-----ATTGAAGATATATGCTTGGCGGACCAAT 198
Qy      49 IleGlyLeuLeuValGlyGlyPheThrMetThrAlaThrTrpValGlyGlyTyrIle 68
Db      199 CTAGTCTGCTGTAATCTGCATTTATCTGACAGGTGCTGATATGTCAGGCTGGCTTTG 258
Qy      69 AsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaTrpAlaHisAlaPro 88
Db      259 CTGGGCTTCCAGGATACATGATACGCTTCAGGCTGTGAATATTGG-----ATCGCA 312
Qy      89 IleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeuPhePheAlaLysProMetAlaArg--- 107
Db      313 CTGGGTCTACCATGCTGCTGTGCTAATTATCTGATTTGTCACGGGCTTGCCTGTT 372
Qy      108 -----SerLysGlyTyrValaIleThrMetLeuAspProPheLysGlnIleTyrGly 123
Db      373 TATACGAGCTTGCAGATATAGCTGTACTTACCAAGATATATTTTCAATGCTTTTCA 432
Qy      124 LysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe----- 140
Db      433 GATTAATCACATCTTTGCTATCATGTCGCGGTGTGCTATATCTATTTTATACGTT 492
Qy      141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValaIleIleAspVala 160
Db      493 TATACAGCAGCACTGTGCTGTGTGTGTAACTGTTGAAGCTCATTTGATCA 552

```

```
OY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
Db 553 TATAGCATGGGTTTGGGTACAGACAGCGCTGTTGGCTTATACCTGTTGGCCGC 612
OY 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200
Db 613 TTTTGGAGGTGTCATTAAGTACCTTTGCAAGCGCTCATTTAGTCAATGCTCAATGCTA 672
OY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db 673 ATT-----GTAACCTGCTGATGCTTGGCGAGATT 702
OY 221 HisAlaLysTyrGlnSerProTyrLeuGlyThrIleGluSerValGlnValTyrThrTyr 240
Db 703 GGTGGTGTCTGACAGCATGCGATGGCCACAGACAAATAACGAGGTCTTTAATGG 762
OY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyr----- 257
Db 763 ATGAATGAC-----GTGACTGTCTCATGGGTGTATCTCTTGTAGTGCATGGGGTTTGGC 816
OY 258 -----PheGlnArgValLeuSerSerSerAla 267
Db 817 TATTTTGGTACCCACACATTTTGTAGCCTTTATGGCAATTCGCTCACTAAAGATGT 876
OY 268 ThrTyrAlaGlnValLeuSer-----PheLeuAlaAlaPheGlyCysLeuVal 283
Db 877 CCGACTGCCATGATGATGGCATGGGCTGATGATTTTAAAGCTTAATGGTCTTGAATG 936
OY 284 MetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyrAsnGlnThr 303
Db 937 GTTGGCTTGGCAGGAGATGCTTATGTGCA-----GCGACA 972
OY 304 AlaTyrGlyTyrProAspProLysThrLysGlnAlaAspMetIleLeuProIleVal 323
Db 973 GGTATAGAGCTTAAGATCTGAAAC-----ATCTTTTGGATATT 1014
OY 324 LeuGlnTyrLeuCysProValTyrIleSerPhePheGlyValAlaValSerAlaAla 343
Db 1015 TCACAGAGTGTCTTCATCCGCTGATTCAGATTTTGTGGCGCAATTTTACAGACA 1074
OY 344 ValMetSerSerPheArgGlnAsnAlaSerAspLysGlnIleValTyrValMetArgIle 363
Db 1075 ATTAAGATGATCATCTCAAGCAATGCTGTGCTCTCAAGTTCACGTGACAGAGATATT 1134
OY 364 TyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTyrValMetArgIle 383
Db 1135 TATAAGTATATTTGGATTAACAAGCTCAGAGCAGCCAGCAAGTGTATGGCCGTATT 1194
OY 384 ThrValIleValPheGlyAlaSerAlaThrAlaMetAla-LeuLeuThrLysThrValTyr 403
Db 1195 TCGGTGTTGTTTATGTCGATCACTTCGATTTGTGGCAGCGCAACAGATTCAAGTA 1254
OY 403 rGlyLeuTyrPyr-----LeuSerSerAspLeuValTyrIleIleIlePheProGlnLe 421
Db 1255 TTAATTTTGATTCATGCTTGGCAGAGATTTGGCGCAGCATTTGGTTCATTG- GTAAT 1313
OY 421 uLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIlePhe 441
Db 1314 TTTAAGCCGTATGTG--AAGCGTATGAACGAAATGTGCTTGGCGTGCATGATTTGT 1370
OY 441 eGly-----LeuPheLeuArgIleThrGlyGlyLysProTyrLeuTyrLeuGlnIle 458
Db 1371 GGGTCCCTTGACGGGATATTGGTGTATGTGT----- 1407
OY 458 cLeuIlePheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsn-----GlnArg 476
Db 1408 -----TTTGAGATTTGGCGGTACAGCCACTAATGATTCATTATTCATTTTACC 1460
OY 476 gPheProPheLysThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyr 496
Db 1461 TTTTCTTT-----AGTTGTGTGACAAACCATTTGAGTAAGTCT 1499
```

```
OY 496 rLeuAlaLysTyrLeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAl 516
Db 1500 GATG-----ACAGCACCCGCGCTGTATTATTTATTTCAAA 1535
OY 516 aValValAlaArgHisSerGluGlnAsnMetAspLys 528
Db 1536 ATTT-----GAGGATATGGAAGAAG 1554

RESULT 10
US-09-596-002-7/c
Sequence 7, Application US/09596002
Patent No. 6632636
GENERAL INFORMATION:
APPLICANT: Lagace, Robert, E.
APPLICANT: Paterson, Chandra
APPLICANT: Berg, Kim, L.
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
FILE REFERENCE: PM-0008-4 US
CURRENT APPLICATION NUMBER: US/09/596,002
PRIOR FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: 60/140,121
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 6703
TYPE: DNA
ORGANISM: M. catarrhalis
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte template ID No. 6632636 7
PUBLICATION INFORMATION:
US-09-596-002-7

Alignment Scores:
Pred. No.: 4,426-23 Length: 6703
Score: 298.00 Matches: 132
Percent Similarity: 41.88 Conservative: 99
Best Local Similarity: 23.94 Mismatches: 223
Query Match: 10.04 Indels: 100
DB: 3 Gaps: 18

US-10-724-806-4 (1-580) x US-09-596-002-7 (1-6703)

OY 9 ValAlaIleIleLeuPheTyrLeuLeuIlePheLeuValGlyIleTyrAlaAlaTyrLys 28
Db 2431 ATTTGGTGGCTGTCTATTATTTGATGATTCGCAATCGCAATTTATGCTTATTTTAA 2372
OY 29 ThrLysAsnSerGlyAsnProGluGlnArgSerGluAlaIleIleValGlyArgAsp 48
Db 2371 CAAAAAAATGAC-----ATTGAAGATATATGCTTGGCGGACGCAAT 2330
OY 49 IleGlyLeuLeuValGlyPheThrMetThrAlaThrTyrValGlyIleTyrIle 68
Db 2329 CTAACTCCCTGCTGATACGATTAATCTGACGTGCTGTGATATCTAGCGCTGTTTG 2270
OY 69 AsnGlyThrAlaGlnAlaValTyrGlyProGlyCysGlyLeuAlaTyrAlaHisAlaPro 88
Db 2269 CTGGGCTTGCCAGATATACATGATACGCTTCAAGCGGTGAAGATTTTG-----ATCGCA 2216
OY 89 IleGlyTyrSerLeuSerLeuIleLeuGlyLeuPhePheAlaLysProMetArg--- 107
Db 2215 CTGGGTTCGACCATTTGGTGTGCTGCTAATTAATCATGATTGTTCACCGCGTTGCGTGT 2156
OY 108 -----SerLysGlyTyrValThrMetLeuAspProPheLysGlnIleTyrGly 123
Db 2155 TATACGAGCTTGCGAGATTAATGCTTACTTATTCACAGATTAATTTTCAATGCTTTTCA 2096
OY 124 LysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyLysMetPhe----- 140
Db 2095 GATAAATCATCACTTTTTCGATCATGTGCGGGGCGGTGATTAATTTTTCACGGTT 2036
OY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
```



```

Db      2035 TATAACGAGCAGCAGTGGTGGTGGTAAAGCTTTGAAAGCTATTGAATTATCA 1976
      :::: |||||
Qy      161 ValaenleleSeValleValSerAlaLeuIleAlaIleLeuTyThrLeuValGlyGly 180
      :::: |||||
Db      1975 TATACATGGGTTTGGGTAAACAGCAGCGCTGGTGGCTTATACCTTGTGGCGGC 1916
      :::: |||||
Qy      181 LeuTySerValAlaTyThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
      :::: |||||
Db      1915 TTTTGGCAGTCTCACTTAACCTGTTGCCAAGGCGTATTATGCGCATGCTA 1856
      :::: |||||
Qy      201 lIeSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
      :::: |||||
Db      1855 ATT-----GTACTGTGGTACACTTTGGCGAGATT 1826
      :::: |||||
Qy      221 HIsAlaIleTyThrGlnSerProTrpLeuGlyThrIleGlnSerValGlnValTyThrTrp 240
      :::: |||||
Db      1825 GGTGGTGTCTGAGGCGATGGCGATTGCCACACAGACAAATACAGAGTCTTTAATTGG 1766
      :::: |||||
Qy      241 LeuAspAsnPheLeuLeuLeuLeuGlyGlyIleProTrpGlnAlaTyThr----- 257
      :::: |||||
Db      1765 ATGAATGCC-----GTGACTGTCACTGGTGTATTCTTTGATGACGAGGGGTTTGGC 1712
      :::: |||||
Qy      258 -----PheGlnArgValLeuSerSerSerSerAla 267
      :::: |||||
Db      1711 TATTTTGGTCACGACACATTAATTGACGCTTATGCGCAATTCGCTCACTTAAGATGTT 1652
      :::: |||||
Qy      268 ThrTyAlaGlnValLeuSer-----PheLeuAlaAlaPheGlyCysLeuVal 283
      :::: |||||
Db      1651 CCGAGCTGCGATGGTGTATGGCATGGCGGATGATTTAAAGCTTAATGGTGTGGTGAAG 1592
      :::: |||||
Qy      284 MetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrpAsnGlnThr 303
      :::: |||||
Db      1591 GTTGGCTTGGCAGGAGTCTTATGTCGA-----CCGACA 1556
      :::: |||||
Qy      304 AlaTyGlyTyProAspProLyThrLyGlnGlnAlaAspMetIleLeuProIleVal 323
      :::: |||||
Db      1555 GGTATAGACCTTAAGATCTGTAACC-----ATCTTTTGGTATT 1514
      :::: |||||
Qy      324 LeuGlnTyThrLeuCysProValTyThrIleSerPhePheGlyLeuGlyAlaValSerAlaIle 343
      :::: |||||
Db      1513 TCACAGAGCTCTTTCATCCGCTGATTTCAGATTCTTGGCGGCAATTTTACGACGA 1454
      :::: |||||
Qy      344 ValMetSerSerAlaSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIle 363
      :::: |||||
Db      1453 ATTAAGAGTACATCTCAAGCAATTCGCTGGTCTCAAGTCTCACTGACACAGATATT 1394
      :::: |||||
Qy      364 TyGlnLeuSerPheArgGlnAsnAlaSerAspLyGlnIleValTrpValMetArgIle 383
      :::: |||||
Db      1393 TATAAGTTATTTTGGATAAACAGCCTCAGAGCGCAGCAAGTCTTATGGCGGTATT 1334
      :::: |||||
Qy      384 ThrValLeuValPheGlyAlaSerAlaThrAlaMetAla-LeuLeuThrLyThrValTy 403
      :::: |||||
Db      1333 TCGGTGTTTGTGTCGATCATTTGGATTATGTCGAGCGCATGCAACAGTTCAGTA 1274
      :::: |||||
Qy      403 rGlyLeuTrpTyThr-----LeuSerSerAspLeuValTyThrIleIlePheProGlnIle 421
      :::: |||||
Db      1273 TTAATTTGGATTCACATGCTTGGCGAGATTTGGCGCAGCATTTGCTCATTTG-CTAAT 1215
      :::: |||||
Qy      421 uLeuCysValLeuPheIleGlyGlyThrAsnThrTyGlyAlaValAlaGlyTyThrIlePhe 441
      :::: |||||
Db      1214 TTTAAGCCGATGTGG--AAGGATATGAACCAATATGGGCTTGGCTGGCATGATTGT 1158
      :::: |||||
Qy      441 eGly-----LeuPheLeuArgIleThrGlyGlyGlnProTyLeuTyLeuGlnTr 458
      :::: |||||
Db      1157 GGTGCTTCGACGCTATTTGGGTGATGCTGT----- 1121
      :::: |||||
Qy      458 oLeuIlePheTyProGlyTyThrTySerAspLyAsnGlyIleTyAsn-----GlnAr 476
      :::: |||||
Db      1120 -----TTGAGATTTGGCGGTGACGACAGTATGATGCAATTATTCACCTTACCGG 1068
      :::: |||||
Qy      476 gPheProPheLyThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTy 496
      :::: |||||

```

```

Db      1067 TTTTGCCTTT-----AGTTGGTGACCAACCATTCAGTAAGTCT 1029
Qy      496 rLeuAlaIleTyThrLeuPheGlnSerGlyThrLeuProProLyLeuAspValPheAspAl 516
      :::: |||||
Db      1028 GATG-----ACACACGCGCGCTGTATTATTTTCACAA 993
      :::: |||||
Qy      516 aValValAlaArgHisSerGlnGlnAsnMetAspLyS 528
      :::: |||||
Db      992 ATTT-----GAGGATATGGAAGAAG 974
      :::: |||||

RESULT 11
US-09-949-016-5012
; Sequence 5012, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5012
; LENGTH: 2450
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5012

Alignment Scores:
Pred. No.: 1,43e-23 Length: 2450
Score: 296.00 Matches: 148
Percent Similarity: 38.1% Conservative: 99
Best Local Similarity: 22.8% Mismatches: 222
Query Match: 9.9% Indels: 180
DB: Gaps: 29

US-10-724-806-4 (1-580) x US-09-949-016-5012 (1-2450)
Qy      11 lIeIleuPheTyThrLeuLeuIlePheLeuValGlyIleTrpAlaAlaTrpLyThrLyS 30
      :::: |||||
Db      104 ATCGTATCTACTCTGCTGTAGTATGCGCGCTGCGCTGTGGCTATGTTTCCACC--- 160
      :::: |||||
Qy      31 AsnSerGlyAsnProGlnGlnArgSerGlnAlaIleIleValGlyGlyArgAspIleGly 50
      :::: |||||
Db      161 AATCGTGGACT-----GTGGAGGCTTCTTCTGCGAGCGCGAAGTATGCG 208
      :::: |||||
Qy      51 LeuLeuValGlyGlyPheThrMetThrAlaThrTrpValGlyGlyTyThrIleAsnGly 70
      :::: |||||
Db      209 TCGTGGCGCATTTGAGCGCTCCCTCTTGTCTAATGATGGAAGTGCACACTTGGGG 268
      :::: |||||
Qy      71 ThrAlaGlnAlaValTyThrGlyProGlyCysGlyLeuAlaTrpAlaHisAlaProIleGly 90
      :::: |||||
Db      269 CTGGCC-----GGGACTGGGAGCGCTTCAGGATCGCCCATTTGAGGC 310
      :::: |||||
Qy      91 TySer-----LeuSerIleIleLeuGlyGlyLeuPhePheAlaLyPro 105
      :::: |||||
Db      311 TTTGAATGAATGCCCTGCTTTTGGTGTGCTGTGGCTGTGCTGTTTTC-----CCG 364
      :::: |||||
Qy      106 MetArgSerLyS--GlyTyValThrMetLeuAspProPheLyGlnIleTyThrLyS 124
      :::: |||||
Db      365 ATCTATATTAAAGCTGGGCTGTGTGCAATGACAGTACTACTGAG-----AAG 412
      :::: |||||
Qy      125 ArgMetGlyGly-----LeuLeuPheIle 132
      :::: |||||
Db      413 CGGTTTGGAGGCGCAGCGGATCCAGGCTTACCTTCCCTTCTGCTGCTGCTTACATT 472
      :::: |||||

```

```

OY 133 ProAlaLeuMetGlyGluMetPheTrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThr 152
Db 473 TTCACCAAGATCTCCGACAGACATCTTCGGGGCCCATATTC----- 514
OY 153 ILeSerValIleIleAspValAspValAsnIleSerValIleValSerAlaLeuIleAla 172
Db 515 ATCAATCTGGCCCTTGGCCCTGATCTGTATTTAGCCATCTTCTCTATTGGCAATCACT 574
OY 173 ILeuTyThrLeuValGlyGlyLeuTySerValAlaTyThrAspValValGlnLeu 192
Db 575 GCCCTTACACAAATTACAGGGGCGGCGGCGGATTTACACGGACCTTGGCAGAG 634
OY 193 PheCyIlePheIleGlyLeuTrpIleSerValProPheAlaLeuSerHisProAlaVal 212
Db 635 GTGATCATGCTGGTGGGCTTTAACTCCTGACTGGGTGTTCTTTCACGAAAGTGGGA--- 691
OY 213 ThrAspIleGlyPheThrAlaValHisAlaIleTyGlnSerProTrpLeuGlyThrIle 232
Db 692 -----GGCTATGAGCCCTTCATGGAAGATCATGAAAGCC---ATTCACACCATTA 739
OY 233 GluSer-----ValGluValTyThr---TrpLeuAspAsn 244
Db 740 GGTCTGTATGCAACACACCTTTCCAGAAAGATCTACATCCAAAGGCCGACCTCTTC 799
OY 245 LeuLeuLeu-----MetLeuGlyGlyIleProTrp----- 254
Db 800 CACATCTTCCGAGATCCCTCCAGGGAGACCTCCACATGGCCCTGGTTCACTTTGGGAGT 859
OY 255 -----GlnAlaTyPheGlnArgValLeuSer 263
Db 860 TCCATCCTTACCTTGTGTGCTGCTGACGACAGATCAGGTATGTGTCAGGCGCTCTTA 919
OY 264 SerSerSerAlaThrTyThrAlaGln-----ValLeuSerPheLeuAlaIleAlaPhe 279
Db 920 GCCAAGAAATATGTCACGTCGAAGGATGGCTGCATCTGTGTGGGATCTTAAGCTGAG 979
OY 280 GlyCySLeuValMetAlaLeuProAlaIle----- 289
Db 980 CCCATGTTTCATCATGTGATGCCAGAAAGATCAGCCGATTCGTACACAGAAAAATT 1039
OY 290 -----CysIleGlyAlaIleGlyAlaSerThrAspTrpAsnGln 302
Db 1040 GCCTGTGCTGCTCCCTTCAGAAATGTGAGAAATTGGGGATCCAAAGTTGGCTGACCAAC 1099
OY 303 ThrAlaTyGlyTyPProAspProLysThrLysGluGluAlaAspMetIleLeuProIle 322
Db 1100 ATCGCCTAT-----CCAAACC 1114
OY 323 ValLeuGlnTyLeuCySerProValTyIleSerPhePheGlyLeuGlyAlaValSerAla 342
Db 1115 TTAGTGTGTGAGCTCATGCGCAATGATGCGAGGCGCTGATGCTATCAGTCACTGCGCC 1174
OY 343 AlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsn 362
Db 1175 TCCCTCATAGCTCCCTGACCTCCATCTTCACACAGGCGCAGACCCCTCTTCACCATGAGC 1234
OY 363 ILeTyGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTrpValMetArg 382
Db 1235 ATCTAC-----GCCAAGGTCGCGCAAGAGAGCATCTGAGAAAGACCTCATGATTCCCGAGAG 1291
OY 383 IleThrValLeuVal---PheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLysThr 401
Db 1292 TTGTTATCTGCTGCTGATTTGGCATCGCATCGCTGGGTCGCCATTTGATGAGTCAGCA 1351
OY 402 ValTyArgLy-----LeuTrpTyLeuSerSerAspLeuValTyThrIle-----IleIle 417
Db 1352 CAAAGTGGGCACTCTGCAATTACATCCAGTCCATACACAGATTACTTGGGACCAACCATT 1411
OY 418 PheProGlnLeuLeuCyValLeuPheIleLysGlyThrAsnThrTyGlyAlaValAla 437
Db 1412 GCGGCTGCTCTCTGCTTGTCTATTTTCTGAAAGAGAGTCAATGAGCCAGGAGCCTTTGG 1471
OY 438 GlyTyRlePheGlyLeuPheLeuArgIle----- 447

```

```

Db 1472 GGACTGATCTCAGACCTTCTGATTGGGATTTACAGTATGATTACTGACTTGGCTTATGGA 1531
OY 448 ThrGly-----GlyGluProTyLeu 454
Db 1532 ACCGGAGCTGCATGAGACCCAGCACTGTCCCAAGATTATCTGTGGGGTGCACACTTG 1591
OY 455 TyLeuGlnProLeuIlePheTyPProGlyTyTyTySerAspLysAsnGlyIleTyAsn 474
Db 1592 TACTTGGCATTAATCTCTTC----- 1612
OY 475 GlnArgPheProPheLysThrLeuSerMetValThrSerPhePheThrAsnIleCyVal 494
Db 1613 -----GCCATTCTTATCATCCATCGATGCTGATC 1642
OY 495 SerTyLeuAlaIleTyLeuPheGluSerGlyThrLeuProProLysLeuAspValPhe 514
Db 1643 TCCCTCCTCACCACCAACCATT-----CCGATGTCGATCTCTAC 1681
OY 515 AspAlaVal-----ValAlaArgHisSerGluGlnAsnMetAspLysThrIleLeuVal 532
Db 1682 CGTCTGTGTGGAGCTCGCAACAGCAAGAGAGCGTATTGAC-----CTGGATCG 1735
OY 533 ArgAsnGlnAsnIleLysLeuAsnGluLeuAlaProValLysProArgGlnSerLeuThr 552
Db 1736 GAAAGGAGAACATCCAAAGAGC-----CTTAAGAGACCATTTAA 1777
OY 553 LeuSerSerThrPheThrAsnLysGlu 561
Db 1778 ATGAAACACAAAGTTCCGAGAGAGAA 1804

RESULT 12
US-09-252-991A-7814/C
; Sequence 7814, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rudenfeld et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7814
; LENGTH: 3342
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1036)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-7814

Alignment Scores:
Pred. No.: 2,396-23 Length: 3342
Score: 296.00 Matches: 123
Percent Similarity: 40.2% Conservative: 96
Best Local Similarity: 22.6% Mismatches: 230
Query Match: 9.9% Indels: 96
DB: 3 Gaps: 16

US-10-724-806-4 (1-580) x US-09-252-991A-7814 (1-3342)
OY 9 ValAlaIleIleLeuPheTyLeuLeuIlePheLeuValGlyIleTrpAla---AlaTrp 27
Db 1638 ATGGCCTCGACATATTGCTGTTCTGATCTAACGCCGGGCGCATGATGCGCTGGGCTGG 1579
OY 28 -----LysThrLysAsnSerGlyAsnProGluGluArgSerGluAla 41

```

Db	1578	TACGGCAATGGCCGGCGGCAAGACCCGGGAGCAC-----	1549
Qy	42	ILleIleValIGlyGlyAAdgAspRIleGlyLeuLeuValIGlyGlyPhe-----ThrMet	58
Db	1545	TACSTGTGGGGCGGGCGGCAACCTGGC-----CCCGGCTTCTACGTGGAAACATG	1495
Qy	59	ThrlAThrTrpValIGlyGlyGlyTyxTrIleAsnGlyThrAlaGluAlaValTyxGlyPro	78
Db	1494	GCSCCAACCGCTCTCGGGGGGCGCTCCACCACTGGCACCTGGCGCTGGGCTAGCTCAT	1435
Qy	79	GlyCyseGlyLeuAlaTrpAlaIleAlaProIIleGlyTyxSerLeuSerLeuIleGly	98
Db	1434	GGCATCTCCGGCTTCTGGCTGTGGCGGCGAGATCGGC-----CTCGCATGTGGCGCTCT	1381
Qy	99	GlyLeuPhePheAlaValProMetArgSerIleGlyTyxValThrMetLeuAspProIhe	118
Db	1380	AGCGTGTCTGGCGGCAACCGCTGTGAAGCTGAAGATCTACACCGTACACCGAGTCTGT	1321
Qy	119	LyseGlnIleTyxGlyTyxArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlu	138
Db	1320	GAGCGTGCTACAAACCCCGCGCGGCGGCAAGCCGCGTATGATCTGTGTACAGCA	1261
Qy	139	MetPheTrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp	158
Db	1260	CTGATGATCGGCGGCAACCTGCACCACTGCACATCGGCACCGTACAGAGTCTGTTCGCG	1201
Qy	159	ValAspValaIleIleSerValIleValSerAlaLeuIleAlaIleIleTyxThrLeuVal	178
Db	1200	CTGCCGTTTGGGTGTGATCTGTATCGGGCGGCGCGCTGTGTCTTACTTCCACATTC	1141
Qy	179	GlyGlyLeuTyxSerValAlaIleTyxThrAspValAlaGlnLeuPheCysIlePheIleGly	198
Db	1140	GGCGGCAATGTGTGTGCTGACCTGCACCACTGTGTCGATTCGATGATACCGTCCGCG	1081
Qy	199	Leu---TrpIleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPhe	217
Db	1080	CTGGTGTCTCGCTGATGCGCGCTGTGCATCAACGACGCGCGC-----GGNTGG	1033
Qy	218	ThrlAlaValAlaIleAlaTyxGlnSerProTrpLeuGlyThrIleGluSerValGluVal	237
Db	1032	GAGCCCGTCGTAAGGCAAGCTGCGGCGGCAAGTACTTC---GACTTCACCGGATGGCGTGG	976
Qy	238	TyrThrTrpLeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyx	257
Db	975	GACACCACTCGCACTACTCTCTCTACTACTCTTCCTTCGGCATCTTCATCGGCAAGACATC	916
Qy	258	PheGlnArGValLeuSerSerSerSerAlaThrTyxAlaGlnValLeuSerPheLeuAla	277
Db	915	TGGCAGCGGGGTCTTCAACCGCGCGGACGAGAACCGGTGGCAAGTGTCCGGCTCCGGCGC	856
Qy	278	AlaPheGlyCysLeuValMetAlaIleProAlaIleCysIleGlyAlaIleGlyAlaSer	297
Db	855	GGTATCTACTCGCTGCTACGCGATGCGCGGCGCGTGAATCGGGATGGCGCAAGGTC	796
Qy	298	ThrlAspTrpAsnGlnThrAlaTyxGlyTyxProAspProIysthIlyGlnGluAlaAsp	317
Db	795	CTG-----CTGCGGAGC-----CTGGAGAAAGCTCAAC	769
Qy	318	MetIleLeuProIleValLeuGlnTyxIleuCysProValTyxIleSerPhePheGlyLeu	337
Db	768	AAGCGCTTCGCAAGTGTGTGAGAGACAGCTGCGGAACGGGATCTCCGGCTGGTATC	709
Qy	338	GlyAlaValaSerAlaAlaValaMetSerSerAlaAspSerSerIleLeuSerAlaSerSer	357
Db	708	GCCCGCCGCGCTGGCGGCGCTAGTATGCCACCGCAAGCGCGGCGTGTGGCGCTGCACG	649
Qy	358	MetPheAlaArgMetIleTyx---GlnLeuSerPheArgGlnAsnAlaSerAspIlyGln	376
Db	648	ACCGTACACCGAGACCTGCGCGCGGCTGTGGCGCGCGCGGCGGCAATCCGACAAAGCG	589
Qy	377	IleValTrpValMetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAla	396
Db	588	GAGGTGACGAGAAACCGATGCGCAACCTGTGCTGGGCGCTGTATGCTCGGATCGCC	529

QY	397	leuLeuThrLysThrValTyrGlyLeuTrpTyrLeuSerSerAspLeuValTyrIleIle	416
Db	528	CTGGTGTGTACAGCAGCATATCAGCCGCCCTGACCCGTGCTTACAACTCTCTGTAGGCGGC	469
QY	417	IlePheProGlnLeuLeuCysValLeuPheIleIleGlyThrAsnThrTyrGlyAlaVal	436
Db	468	ATGCTGATCCCGCATCGATCGGCGCATCTACATGGAAGCGCGCACCAACCCGCGCGCATTC	409
QY	437	AlaGlyTyrIlePheGlyLeuPheLeuAsnArgIleThrGlyGlyGlnProTyrLeuTyrLeu	456
Db	408	ACCAAGCATGACCCCTGGGC-----TTCCTT	385
QY	457	GlnProLeuIlePheTyr-----ProGlyTyrTyrSer	467
Db	384	ACCGTACTCGGTTCATGATCATGAAGAGCGGCTTCGACGCCAATAGCGCATCTACTACAC	325
QY	467	-----	467
Db	324	CTGGCGGTGGCCTGCTCAGCTTCGTCTGTCAGCTTCTTCGCGGCGCCGACGAGT	265
QY	468	-----AspLysAsnGlyIleTyrAsnGlnArg	476
Db	264	GTCGCCAGCGCGCCTGAAACCAAGACACAACCCGACAGCGAGGAACTTTCCTCGCCTT	205
QY	477	PheProPheLysThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyr	496
Db	204	TTCCATTCCCGAAGACATGAATAAAGAACGCGCATGACAC-----TGCGGCGCAATTC	151
QY	497	LeuAlaLysTyrLeu	501
Db	150	CTGCTCAAGCAACTC	136

```

RESULT 13
US-09-252-991A-7678
; Sequence 7678, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7678
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (604)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-7678

```

Alignment scores:	
Pred. No.:	1,18e-23
Score:	293.50
Percent Similarity:	42.5%
Best Local Similarity:	23.8%
Query Match:	9.8%
DB:	3
Length:	1455
Matches:	114
Conservative:	90
Mismatches:	213
Indels:	63
Gaps:	14

US-10-724-806-4 (1-580) X US-09-252-991A-7678 (1-1455)

```

QY      9 ValaIaIeIleLeuPheTyrLeuLeuIlePheLeuValGlyIleTrrpala--AlaTrp 27
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      2 ATGCGCCCTGCACATATTCGTCTTCTATCTCAAGCCGCGGCGACATCGGCGCTGGCTGG 61

```

```

QY 28 -----LysThrIysAsnSerGlyAsnProGluGluArgSerGluAla 41
Db 62 TACGGCATGCGCGCGCGCAAGACCGCGACGAC----- 94
QY 42 IleIleValGlyGlyArgAspIleGlyLeuValGlyGlyPhe-----ThrMet 58
Db 95 TACCTGGTGGCGCGCGCACTCGGC-----CCGGGCTTCACTCGGGAACCATG 145
QY 59 ThrAlaThrTrpValGlyGlyGlyTrpIleAsnGlyThrAlaGluAlaValGlyPro 78
Db 146 GCCGCGACCGCTCTCGCGCGCGCTCCACCATCGGCGCGCTGGGCTGACCTCAT 205
QY 79 GlyCysGlyLeuAlaTrpAlaHisAlaProIleGlyTrpSerLeuSerLeuIleGly 98
Db 206 GGCATCTCGCGCTTCGCTGCGCGCGATCGGC-----CTCGGCATCTCGCGCTC 259
QY 99 GlyLeuPhePheAlaIysProMetArgSerIysGlyTrpValThrMetLeuAspProPhe 118
Db 260 AGCTGTTCTGCGCGCGCGCTGCTGAAGCTGAAGATCAACGCTCACCGAGTCTCG 319
QY 119 LysGlnIleTrpGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlu 138
Db 320 GAGCGTTCCTAACACCGCGCGCGCGCACCGCGCGCTGATCATGCTGCTACAGCA 379
QY 139 MetPheTrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp 158
Db 380 CTGATGATCGGCGCGCACCTCGACCATCGCATCGGCAACCGTATCAGGCTCTGTCGCGC 439
QY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuThrTrpVal 178
Db 440 CTGCGCTTCTGGGTCTCATCTGATCGCGCGCGCGCTGCTGCTCTACTCCACCATC 499
QY 179 GlyGlyLeuTrpSerValAlaTrpThrAspValAlaGlnLeuPheCysIlePheIleGly 198
Db 500 GCGCGCATAGTGGTCCGCTGACCTGACCGCATCGTGCAGTTCCTGATCATGACCGTCCGCGC 559
QY 199 Leu---TrpIleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPhe 217
Db 560 CTGCTGTTCTCGCTGATGCGCTGCTGATCAACGACGCGCGC-----GAGTGG 607
QY 218 ThrAlaValHisAlaIleTrpGlnSerProTrpLeuGlyThrIleGlySerValGluVal 237
Db 608 GACGCGCTCGTGAAGCCAGCTGCGCGCGCGCTACTTC---GACTTCAACCGCATCGGCTGG 664
QY 238 TyrThrTrpLeuAspAsnPheLeuLeuLeuMetLeuGlyIleProTrpGlnAlaTrp 257
Db 665 GACACCATCTGCATCTACTTCTGATCTACTTCTCGGATCTTCATCGCGCAGACATC 724
QY 258 PheGlnArgValAlaLeuSerSerSerAlaThrTrpAlaGlnValLeuSerPheLeuAla 277
Db 725 TGGCAGCGGAGTTCACCGCGCGCGCGACGCTGGCCACGCTGCGCGCTCCGCGCGC 784
QY 278 AlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSer 297
Db 785 GGTATCTACTGCTGCTCTACGCGCATGCGCGGGGCGTGAATGGGATGGCGCGCAAGCTC 844
QY 298 ThrAspTrpAsnGlnThrAlaTrpGlyTrpProAspProLysThrLysGluGluAlaAsp 317
Db 845 CTG-----CTGCGCGGAC-----CTGCGAGAACCTGCATC 871
QY 318 MetIleLeuProIleValLeuGlnTrpLeuCysProValTrpIleSerPheGlyLeu 337
Db 872 AACGCTTCGCCAGTGGTGCAGACACGCTGCGAGACGCGATCGCGGCTGTGTCATC 931
QY 338 GlyAlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSer 357
Db 932 GCCGCGCGCGCTGGCGCGCTGATGTCACCGCGCAGCGCGCGCTGCGCGCGCTGCAC 991
QY 358 MetPheAlaIleAsnIleTrp---GlnLeuSerPheArgGlnAsnAlaSerLysGlu 376
Db 992 ACCGTACACGAGACCTGCTGCGCGCGCTGCGCGCGCGCGCGCGCAACGCTCCACGAGC 1051
QY 377 IleValTrpValMetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAla 396

```

```

Db 1052 GACGCGACGAGAACCGCATCGCGACCGCTGCTGCGCGCTGATAGTCTCGCATTCGC 1111
QY 397 LeuLeuThrLysThrValTrpGlyLeuTrpTrpLeuSerSerAspLeuValTrpIleIle 416
Db 1112 CTGCTGCTCAGCAGCATATATACGCGCGCTGACCGTGGCTTACACCTCTGCTAGCGCGC 1171
QY 417 IlePheProGlnLeuLeuCysValLeuPheIleLysGlyThrAsnThrTrpGlyAlaVal 436
Db 1172 ATGCTGATCCCGCTGATGCGCGCATCTGAAGCGCGCACCGCGCGCGCGCATC 1231
QY 437 AlaGlyTrpIlePheGlyLeuPheLeuArgIleThrGlyGlyGluProTrpLeuTrpLeu 456
Db 1232 ACCAGCATGACCTCGGC-----ProGlyTrpTrpSer 467
QY 457 GlnProLeuIlePheTrp-----ProGlyTrpTrpSer 467
Db 1256 ACCGACTGCTGCTTCAATATCAAGACGCGCTCGACGCCATACCGCATCTACTACAGC 1315

RESULT 14
US-09-252-991A-7528
; Sequence 7528, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7528
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (696)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-7528

Alignment Scores:
Pred. No.: 1,22e-23 Length: 1485
Score: 293.50 Matches: 114
Percent Similarity: 42.5% Conservative: 90
Best Local Similarity: 23.8% Mismatches: 213
Query Match: 9.8% Indels: 63
DB: 3 Gaps: 14

US-10-724-806-4 (1-580) x US-09-252-991A-7528 (1-1485)
QY 9 ValAlaIleIleLeuPheTrpLeuLeuIlePheLeuValGlyIleTrpAla---AlaTrp 27
Db 94 ATGCGCTTCGACATATTCGCTGCTGATCTGACCGCGGAGCATGATGCGCTGGGCTGG 153
QY 28 -----LysThrIysAsnSerGlyAsnProGluGluArgSerGluAla 41
Db 154 TACGGCATGCGCGCGCGCAAGACCGCGACGAC----- 186
QY 42 IleIleValGlyGlyArgAspIleGlyLeuValGlyGlyPhe-----ThrMet 58
Db 187 TACCTGGTGGCGCGCGCACTCGGC-----CCGGGCTTCACTCGGGAACCATG 237
QY 59 ThrAlaThrTrpValGlyGlyGlyTrpIleAsnGlyThrAlaGluAlaValGlyPro 78
Db 238 GCCGCGACCGCTCTCGCGCGCGCTCCACCATCGGCGCGCTGGGCTGACCTCAT 297
QY 79 GlyCysGlyLeuAlaTrpAlaHisAlaProIleGlyTrpSerLeuSerLeuIleGly 98

```

Db	298	GGCATCTCCGGCTTCTGGCTGTGGGGCGCATCGGC-----CTGGCATGCTGGGGCTC	351
Qy	99	GlyLeuPhePheAlaIalYProMetArgSerIysGlyTyrValThrMetLeuAspProPhe	118
Db	352	AGCGTCTTCCTGGCGCAAGCCGCGCTCTGAGCTGAAGATCTACACCAACCCAGCTCTG	411
Qy	119	LysGlnIleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlu	138
Db	412	GAGCGTGGCTTCAACCCCGCGCGCGCCAGCCAGCGCCCTGATCATGCTGATCTACGCA	471
Qy	159	ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuVal	178
Db	532	CTGCCGTTCTTGGCTCGATCTGATCTGATCGCGCGCGCGCTGCGTCTCTACCAACATC	591
Qy	179	GlyGlyLeuTyrSerAlaIalTyrThrAspAlaValGlnLeuPheCysIlePheIleGly	198
Db	592	GCGCGCATGTGGTGGCTGACCTGACCCCTGACCAATCGTGGATTCTGATCATACCGTGGC	651
Qy	199	Leu---TrpIleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPhe	217
Db	652	CTGGTGTCTTCCTGCTGATGATCCGCTGTGCATCAACAGCCGGC-----GAGTGG	699
Qy	218	ThrAlaValHisAlaIalYsTyrGlnSerProTrpLeuGlyThrIleGluSerValGluVal	237
Db	700	GAGCCCTCGTGAAGCCAGCTCCGCGCCAGCATATCTC---GACTTCACCGCGATGGCGTG	756
Qy	238	TyrThrTrpLeuAspAsnPheLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyr	257
Db	757	GACACCAATCGGCACCTACTCTCTCATCTTCTTCGGCATCTTCATCGCCAGACATC	816
Qy	258	PheGlnTrpValIleuSerSerSerAlaThrTyrAlaGlnValIleuSerPheLeuAla	277
Db	817	TGGCAGCGGGGTTCACCCGCCGACGACCCGTGGCAAGTGGCCGCTCCGCGCC	876
Qy	278	AlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSer	297
Db	877	GGTATCTACTCGGCTCTACGGCATGGCGGGGCGCTGATCGGATGGCCGCAAGTTC	936
Qy	298	ThrAspTrpAsnGlnThrAlaIalTyrGlyTyrProAspProLysTrpLysGluGluAlaAsp	317
Db	937	CTG-----CTGCCGAGAC-----CTGGAGAACGTCAAC	963
Qy	318	MetIleLeuProIleValLeuGlnTyrIleuCysProValTyrIleSerPhePheGlyLeu	337
Db	964	AACGCTTCCTCGCAGCTGGTGGAGACAGCTCTCGGAACGGCATTCGCGGCTGGTCACT	1023
Qy	338	GlyAlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSer	357
Db	1024	GCCGCGCGCGCTGGCGCGCTATGTCCACCGCCAGCGCGCGCTGCTGGCGGCTCCAC	1083
Qy	358	MetPheAlaArgAsnIleTyr---GlnLeuSerPheArgGlnAsnAlaSerAspLysGlu	376
Db	1084	ACCCTGACCCGACAGACCTGCTCGCGGCTCTGGCGCGCGCGCGCCAGATCCGACAAACGC	1143
Qy	377	IleValTrpValMetGlyIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAla	396
Db	1144	GACGTGACGAGAACCGCATGCGCAGCCTGTGCTGGCGCTGTGATGCTCGGATGCTCC	1203
Qy	397	LeuLeuThrLysTrpValTyrGlyLeuTrpTyrLeuSerSerAspLeuValTyrIleIle	416
Db	1204	CTGTGTGTGACGACGTGATCAGCGCCCTGACCGGTGCTACACCTGCTGTAGCGCGC	1263
Qy	417	IlePheProGlnLeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaVal	436
Db	1264	ATGCTGATCCGCGTATATCGGCGCATCTACTGAGAAAGCCGACACACCGCGCGGCTATC	1323
Qy	437	AlaGlyTyrIlePheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeu	456
Db	1324	ACCAAGCATGACCTGGGC-----TTCCTT	1347

```

QY      457 GlnProLeuIlePheTyT-----ProGlyTyTYrSer    467
Db      1348 ACCGACTGGGTTCATCATCAGACGCCTCCGACGCCAATACCGCCATTCTACTACAGC    1407

RESULT 15
US-07-841-651-1
; Sequence 1, Application US/07841651
; Patent No. 5410031
; GENERAL INFORMATION:
; APPLICANT: Fajor, Ana M
; APPLICANT: Wright, Ernest M
; TITLE OF INVENTION: Cloning and Functional Expression of a
; TITLE OF INVENTION: Mammalian Na+/Nucleoside Cotransporter: A Member of the
; TITLE OF INVENTION: SGLT Family
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/841,651
; FILING DATE: 19920224
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandel, Saralynn
; REGISTRATION NUMBER: 31,853
; REFERENCE/DOCKET NUMBER: 8772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2238 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..2022
; US-07-841-651-1

Alignment Scores:
Pred. No.:          3,05e-22           Length:         2238
Score:              284.00             Matches:         150
Percent Similarity: 36.2%               Conservative:     82
Best Local Similarity: 23.4%            Mismatches:     226
Query Match:        9.5%                Indels:         182
DB:                 2                   Gaps:          26

US-10-724-806-4 (1-580) x US-07-841-651-1 (1-2238)

QY      8 LeuValAlaIleIleLeuPheTyTLeuLeuIlePheLeuValGlyTLerrPaLaLaTaTP    27
Db      82 ATCGGGTCATVCTGCTCTTAATTTCTCGTGTCATTTGGTGTGGCTTGTGGTCATGTGNC    141

28 LysThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyGlyArg    47
Db      142 AGAACCC---AACAGAGCACCC-----GTGGGTGGCTACTTCTCTGGCAGAGACA    186

```

